

Improved Error Bounds for the Adiabatic Approximation

Donny Cheung,¹ Peter Høyer,¹ and Nathan Wiebe¹

¹*Institute for Quantum Information Science, University of Calgary, Alberta, Canada*

Since the discovery of adiabatic quantum computing, a need has arisen for rigorously proven bounds for the error in the adiabatic approximation. We present in this paper, a rigorous and elementary derivation of upper and lower bounds on the error incurred from using the adiabatic approximation for quantum systems. Our bounds are often asymptotically tight in the limit of slow evolution for fixed Hamiltonians, and are used to provide sufficient conditions for the application of the adiabatic approximation. We show that our sufficiency criteria exclude the Marzlin–Sanders counterexample from the class of Hamiltonians that obey the adiabatic approximation. Finally, we demonstrate the existence of classes of Hamiltonians that resemble the Marzlin–Sanders counterexample Hamiltonian, but also obey the adiabatic approximation.

I. INTRODUCTION

The adiabatic approximation is a technique of central importance to the study of time-dependent quantum systems. Its importance is derived from the fact that adiabatic evolution is one of the few known limiting cases where the dynamics that is generated by a time-dependent Hamiltonian is well understood. The approximation states that if the Hamiltonian that generates a quantum system’s evolution changes sufficiently slowly, then a quantum state that is prepared in an instantaneous eigenstate remains in an instantaneous eigenstate throughout the evolution. The adiabatic approximation has been used to control the evolution of quantum systems [1, 2], analyze particle collisions [3], design circuit based quantum computers [4–6] and has been used to create a new paradigm for quantum computing; namely adiabatic quantum computing [7, 8].

In many of these applications, bounding the time needed to apply the adiabatic approximation was seen as a secondary issue. Adiabatic quantum computing proved to be a notable exception. In adiabatic quantum computing, quantum algorithms are constructed by choosing a time-dependent Hamiltonian whose ground state at $t = 0$ is an easily prepared quantum state and at $t = T$, the ground state encodes the solution to a computational problem. If the adiabatic approximation holds, then the time-evolution transforms the initial ground state into the solution to the problem, thereby enacting the quantum algorithm. A different notion of algorithmic complexity must be used to assess these algorithms because they do not use a discrete set of gates. The scaling of the energy and the time needed for this evolution with the problem size, is a natural way to assess the complexity of such algorithms [7, 9]. Consequently, rigorous estimates for the error in the adiabatic approximation are needed to demonstrate such a scaling [10].

There are many possible ways to adiabatically evolve the initial state into the final state [11–13]. We are, at least in principle, given the freedom to choose these Hamiltonians to depend on the properties of the evolution such as the Hilbert space dimension, the error tolerance or even the total evolution time. This line of inquiry lead to the Marzlin–Sanders counterexample [14], which overturned conventional wisdom about the validity of the approximation.

In 2004 Marzlin and Sanders [14] put forward a Hamiltonian that satisfied the traditionally held “slowness” criterion for the validity of the approximation, and yet it generated an evolution that did not agree with the adiabatic approximation, even in the limit of arbitrarily slow evolution. This created doubt in the scientific community about the conditions that are necessary for the adiabatic approximation and, along with the discovery of adiabatic quantum computing, has led to further interest into the conditions are needed to apply the adiabatic approximation [15–17].

We achieve three important goals in this paper. First, we provide a rigorous analysis of adiabatic evolution that is intuitive, elementary and avoids ad hoc physical assumptions. Second, we provide upper bounds and lower bounds for the error in the adiabatic approximation that are often asymptotically tight, in the limit of slow evolution for a fixed Hamiltonian. Third and last, we show that the adiabatic approximation can be successfully applied to a broad class of quantum systems that are related to the Marzlin–Sanders counterexample.

II. BACKGROUND

In this section we present the background material that is needed to understand our main results, which are given in the following section. In particular we discuss the adiabatic approximation in greater detail, the Marzlin–Sanders counterexample [14] and the Jansen, Ruskai and Seiler (JRS) bound on the error in the adiabatic approximation [10].

The time-evolution of closed quantum systems over a time interval $t \in [0, T]$ is described by the time-evolution

operator, which is expressed in terms of the dimensionless time $s = t/T$ as

$$U(s, 0)|\psi(0)\rangle := |\psi(s)\rangle, \quad (1)$$

where $|\psi(s)\rangle$ is a quantum state evaluated at the dimensionless time s .

We write the adiabatic approximation for the action of the time-evolution operator on an instantaneous eigenstate of $H(0)$, $|G(0)\rangle$ (typically taken to be the ground state) as

$$U(s, 0)|G(0)\rangle \approx e^{-i \int_0^s E_G(\xi) d\xi T} |G(s)\rangle, \quad (2)$$

where $E_G(s)$ is the eigenvalue of the instantaneous eigenvector $|G(s)\rangle$, and the geometric phase [18] is absorbed into the definition of the eigenstate. The error in the adiabatic approximation is the magnitude of the projection of the evolved state onto the space that is orthogonal to $|G(s)\rangle$. This error is generally considered to be small if the variation of the Hamiltonian with time is small compared to the square of the minimum eigenvalue gap between the initial state $|G(s)\rangle$, and all other instantaneous eigenvectors [19, 20]. We write this requirement as

$$\frac{\max_s \|\dot{H}(s)\|}{\min_{s,\nu} [\gamma_{\nu,G}(s)]^2} \in o(T), \quad (3)$$

where $\gamma_{\nu,G}(s)$ is the eigenvalue gap between the instantaneous eigenstates $|\nu(s)\rangle$ and $|G(s)\rangle$. Adiabatic theorems have also been proven for systems that have eigenvalue crossings, but approach those crossings in a smooth fashion [21].

The Marzlin–Sanders counterexample [14] provided a Hamiltonian for which Eq. 3 holds but Eq. 2 does not. Their counterexample Hamiltonian is chosen such that $\|\ddot{H}(s)\|$ diverges in the limit as the quantum system’s evolution becomes arbitrarily slow. The discrepancy with Eq. 2 was attributed to a resonant effect, wherein the normally negligible contribution of the second derivative of H became dominant due to increasingly rapid variation as T increases. Although the fact that the error in the adiabatic approximation depends on the differentiability class of the Hamiltonian was recognized before the Marzlin–Sanders counterexample was discovered [22–24], the counterexample showed that terms that depend on higher-order derivatives of $H(s)$ can sometimes dominate the expression for the error in the adiabatic approximation.

This problem can be addressed by using upper bounds for the error in the adiabatic approximation, rather than asymptotic estimates, because upper bounds do not neglect any contribution to the error. Examples of such upper bounds can be found in [10, 23–25]. The best upper bound presently known for the error in the adiabatic approximation is given by the JRS bound [10]. The JRS bound states that for an m -fold degenerate initial state $|G(0)\rangle$ (meaning that there exist $m - 1$ instantaneous eigenvectors that share the same eigenvalue as $|G(s)\rangle$ for all s), the approximation error $\|(\mathbb{1} - P_G(1))U(1, 0)|G(0)\rangle\|$ is bounded above by

$$\frac{m\|\dot{H}(0)\|}{\gamma(0)^2 T} + \frac{m\|\dot{H}(1)\|}{\gamma(1)^2 T} + \frac{1}{T} \int_0^1 \left(\frac{m\|\ddot{H}(\xi)\|}{\gamma(\xi)^2} + \frac{7m\sqrt{m}\|\dot{H}(\xi)\|^2}{\gamma(\xi)^3} \right) d\xi. \quad (4)$$

The quantity $\gamma(\xi)$ in Eq. 4 represents the minimum eigenvalue gap between the state $|G(\xi)\rangle$ and any other instantaneous eigenstate at time $s = \xi$, $P_G(1)$ represents the projector $|G(1)\rangle\langle G(1)|$ and $\mathbb{1}$ is the identity operator. Given that the Hamiltonian $H(s)$, is not explicitly a function of T , this upper bound can be inverted to find a value of T such that the error in the adiabatic approximation is less than any fixed error tolerance.

As Eq. 4 depends on $\|\ddot{H}(s)\|$, it seems to be ideal for addressing the issues caused by the divergent second derivative of the Marzlin–Sanders counterexample Hamiltonian. However, JRS noted that the third term in Eq. 4 can be replaced by one that is $O(1/T^2)$ [10, 23], implying that their upper bound is not asymptotically tight. Because their bound is not asymptotically tight, it cannot be used to find a sufficient condition for the validity of Eq. 3.

We expand on these results by providing upper and lower bounds for the error in the adiabatic approximation that are often asymptotically tight (meaning that our error bounds converge to the observed error in the limit of large T), given that the leading order term in our bounds does not vanish for any T . We then use these bounds to address Hamiltonians that are related to the Marzlin–Sanders counterexample. We find using these bounds that many Hamiltonians whose second derivatives diverge as T increases, obey the adiabatic approximation. These results are summarized in the following section.

III. RESULTS

This section contains our main results, which address the question of when the adiabatic approximation can be used and provide asymptotic estimates of the error in the approximation. These results are consequences of our upper bounds for the error in the adiabatic approximation, which we present in Section VII.

As with most adiabatic theorems, we require that the variation of the Hamiltonian is slow compared to an appropriate timescale. Our first theorem states that the adiabatic approximation is valid if the evolution time T , is large compared to a timescale Δ_0 , which we define below. Our timescale Δ_0 reduces to that in Eq. 3 if T is large compared to a second timescale Δ_1 , which depends on the second and third derivatives of the Hamiltonian. These timescales are given in the following definition.

Definition 1. Let $H(s)$ be differentiable three times on the interval $[0, 1]$, $\gamma_{\nu, G}(s)$ be the eigenvalue gap between the instantaneous eigenstates $|G(s)\rangle$ and $|\nu(s)\rangle$ and let γ_{\min} be the minimum eigenvalue gap between any two non-degenerate instantaneous eigenvectors of $H(s)$ on this interval. We then define the timescales Δ_0 and Δ_1 to be

$$\Delta_0 := \frac{\|\dot{H}\|}{\min_{\nu, s} [\gamma_{\nu, G}(s)]^2} + \frac{\Delta_1}{\gamma_{\min} T}, \quad (5)$$

$$\Delta_1 := \frac{1}{\gamma_{\min}} \left(\frac{\|\ddot{H}\|}{\gamma_{\min}} + \frac{\|\ddot{H}\|^3}{\gamma_{\min}^3} + \frac{\|\dot{H}\|^6}{\gamma_{\min}^6} \right), \quad (6)$$

where $\|\dot{H}\|$, $\|\ddot{H}\|$ and $\|\ddot{H}\|$ represent the maximum values of the norms of the first three derivatives of $H(s)$.

The following theorem uses the criterion that the timescale Δ_0 is asymptotically smaller than T as a sufficient condition for the applicability of the adiabatic approximation, in the limit of large T . We refer to the resulting theorem as a zeroth-order adiabatic theorem because it gives an expression for the approximation error that is correct to zeroth-order in powers of T^{-1} .

Theorem 1. (Zeroth-order adiabatic theorem) Let $H : [0, 1] \mapsto \mathbb{C}^{N \times N}$ be a Hamiltonian that is differentiable three times and has a minimum eigenvalue gap between any two non-degenerate eigenvectors of $\gamma_{\min} > 0$. If $|G(s)\rangle$ is a non-degenerate instantaneous eigenstate of $H(s)$ and $\Delta_0 \in o(T)$, then the error in the adiabatic approximation obeys

$$(\mathbb{1} - P_G(1))U(1, 0)|G(0)\rangle = O(\Delta_0/T), \quad (7)$$

where $\mathbb{1}$ is the identity operator.

The above theorem states that the error in the adiabatic approximation is zero, plus an additional term that is small in the limit of large T . The requirement that $\Delta_0 \in o(T)$, reduces to the standard criterion given in Eq. 3 if the derivatives of $H(s)$ are independent of T . This implies that Theorem 1 is as a generalization of the standard criterion for the validity of the approximation. The following theorem improves upon the result of Theorem 1 by providing an expression for the error in the adiabatic approximation that is correct to leading order in powers of T^{-1} . We call this result a first-order adiabatic theorem.

Theorem 2. (First-order adiabatic theorem) Let $H : [0, 1] \mapsto \mathbb{C}^{N \times N}$ be a Hamiltonian that is differentiable three times and has a minimum eigenvalue gap between any two non-degenerate eigenvectors of $\gamma_{\min} > 0$. If we take $|G(s)\rangle$ to be a non-degenerate instantaneous eigenstate of $H(s)$ and $\Delta_1 \in o(\gamma_{\min} T^2)$, then the error in the adiabatic approximation, $(\mathbb{1} - P_G(1))U(1, 0)|G(0)\rangle$, is

$$\sum_{\nu \neq G} e^{-i\phi_\nu} |\nu(1)\rangle \left(\frac{\langle \dot{\nu}(s_1) | G(s_1) \rangle e^{-i \int_0^{s_1} \gamma_{G, \nu}(\xi) d\xi T}}{-i \gamma_{G, \nu}(s_1) T} \right) \Bigg|_{s_1=0}^1 + O\left(\frac{\Delta_1}{\gamma_{\min} T^2}\right), \quad (8)$$

where $\phi_\nu = \int_0^1 E_\nu(\xi) d\xi T$, $\gamma_{G, \nu}(s) := E_G(s) - E_\nu(s)$, $P_G(1) = |G(1)\rangle \langle G(1)|$ and the phase of each instantaneous eigenstate is chosen such that $\langle \dot{\nu}(s) | \nu(s) \rangle = 0$ for every $s \in [0, 1]$ and every $\nu \in \{0, \dots, N-1\}$.

It may seem strange that the leading order expression in Eq. 8 only depends on the properties of the Hamiltonian at the beginning and end of the adiabatic evolution, in contrast to Eq. 3. This dependence occurs because the time-evolution operator is expressed as a sum of rapidly oscillating path-integrals in the adiabatic limit [26]. We expect that the contribution of each complete oscillation will have a negligible contribution to the integral if the oscillation frequency is sufficiently rapid compared to the timescale for the variation of the Hamiltonian. Therefore we expect that the contribution of the *incomplete* oscillations will dominate the behavior of the integral. Only the first and last such oscillations are not necessarily complete. The contribution of the first the last and last oscillation is dominated by the properties of the Hamiltonian at the beginning and the end of the evolution respectively; therefore, we expect that the leading order expression for the error in the adiabatic approximation should be dictated by the properties of the Hamiltonian at those times. This dependence is also noted by a number of other studies [10, 13, 26, 27].

These theorems imply that the adiabatic approximation also applies for some Hamiltonians that behave like the Marzlin–Sanders counterexample in that $\|\dot{H}\|$ and $\|\ddot{H}\|$ are increasing functions of T . Our first corollary, which is given below, provides a criterion for the size of the derivatives of H that is sufficient to guarantee that the standard estimate of the error in the adiabatic approximation applies.

Corollary 1. *If $H : [0, 1] \mapsto \mathbb{C}^{N \times N}$ is differentiable three times, $|G(s)\rangle$ is a non-degenerate eigenstate of $H(s)$, $\|\ddot{H}\| \in o(\sqrt{T})$, $\|\ddot{H}\| \in o(T)$ and both $\|\dot{H}\|$ and $(\gamma_{\min})^{-1}$ are bounded above by a constant function of T for all $T \geq 0$ then*

$$\|(\mathbb{I} - P_G(1))U(1, 0)|G(0)\rangle\| \in O\left(\max_{s=0,1} \left[\frac{\|\dot{H}(s)\|}{\min_{\nu} [\gamma_{\nu,G}(s)]^2 T} \right]\right). \quad (9)$$

In other cases the adiabatic approximation may hold, but the error in the adiabatic approximation may not be $O(1/T)$. A sufficient condition for the validity of the adiabatic approximation, for a sufficiently large T , is given below.

Corollary 2. *If $H : [0, 1] \mapsto \mathbb{C}^{N \times N}$ is twice differentiable, $|G(s)\rangle$ is a non-degenerate eigenstate of $H(s)$, $\|\ddot{H}\| \in o(T)$ and both $\|\dot{H}(s)\|$ and $(\gamma_{\min})^{-1}$ are bounded above by a constant function of T for all $T \geq 0$ then*

$$\lim_{T \rightarrow \infty} \|(\mathbb{I} - P_G(1))U(1, 0)|G(0)\rangle\| = 0. \quad (10)$$

A principal limitation of these results is that we take $|G(s)\rangle$ to be non-degenerate for all s . We make this assumption to simplify our analysis of the time-evolution operator, and leave the general case for future work.

These results are laid out as follows: in Sec. IV we discuss the concept of paths in greater detail and their relation to time-evolution. In Sections V and VI we use the results of Sec. IV to provide a path integral representation of the time-evolution operator. In Sec. VII we use integration by parts to approximate the path integrals and prove Theorems 1 and 2. We then compare our error bounds to the error found for the Search Hamiltonian in Sec. VIII and discuss the Marzlin–Sanders counterexample [14] and prove Corollaries 2 and 1 in Sec. IX.

IV. PATHS

The concept of paths was first introduced by Wiener in 1924 [28], in which he showed that the probability distribution for a classical system is found by summing the final states of all paths that could describe the evolution of the system. The corresponding result for quantum systems was proven by Feynman in 1948 [29], and a version that uses the Hamiltonian, rather than the Lagrangian, is given by Farhi and Guttman [30]. The latter approach is used by Mackenzie, Marcotte and Paquette (MMP) [26] to find an expression for the time-evolution of adiabatic quantum systems using perturbation theory. We provide the background material that is needed to understand what we mean by paths in this section, and in the subsequent section, we use these results to derive the MMP representation of the time-evolution operator, but without using perturbation theory.

Our usage of path integrals is somewhat different than Feynman’s original formulation. We obtain our path integral representation, similarly to Feynman, by first discretizing time. By approximating the Hamiltonian with one that is piecewise constant on these discrete time steps, we obtain an approximate evolution operator that is a product of exponentials of the Hamiltonian evaluated at each of the discrete times. Unlike Feynman, we switch to the instantaneous eigenbasis of the Hamiltonian at the given time-step. The resulting expression is then simplified by expanding the product of sums into a sum of products of the projectors that are used to represent the basis transformations. Each product of projectors that occurs in this expansion can then be interpreted as a description of a *potential* evolution of the system, which we call a path. We then take the limit as the number of time-steps in our discrete time-evolution operator approaches infinity. This removes the approximation error caused by taking the Hamiltonian to be piecewise constant and turns the sum over all such paths into a sum of path integrals. These ideas will be elaborated on as we proceed.

We now present a more formal discussion what we mean by a path. A path is a description of a sequence of states that could describe a sequence of observations of the evolution of the quantum state. In particular, we take a path to be an ordered pair $(\{\nu\}, \{s\})$, where $\{\nu\}$ is a list of labels that describe the instantaneous eigenstates of $H(s)$ and $\{s\}$ specifies the time that the system transitions to that state.

We require that no two consecutive labels in a path are identical, ie., $\nu_q \neq \nu_{q+1}$ for all q . With this restriction, it is clear that the elements of $\{s\}$ correspond to the times at which the system jumps from one state to another. We also take the sequence s to be a strictly increasing sequence of dimensionless times taken from the interval $[0, 1]$. We

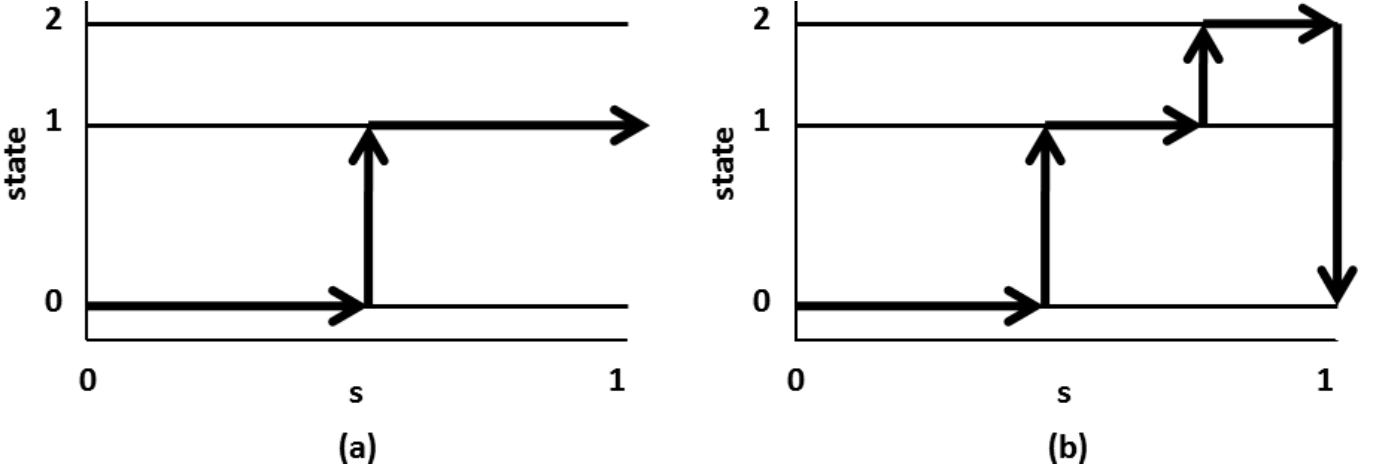


FIG. 1: These diagrams are a representation of the paths used in the examples in this section. Here the state labels refer to different values of ν and do not necessarily refer to energy.

take these times to be strictly increasing because paths that move backwards in time, or are in two different states at the same time, cannot describe a physically realistic sequence of measurements of the state in its instantaneous eigenbasis. As a final note, we first consider the case where each path is composed of at most $L + 1$ distinct states and times, where the times in $\{s\}$ are taken to be integer multiples of $1/L$ to simplify the analysis, similarly to Ambainis and Regev [25]. Unlike the formulation of their result, we take the limit as L approaches infinity early in our analysis thereby removing L from our final results.

As an example of a path, consider the evolution for a three-state system with $L = 4$ given in Fig. 1a: $|0(0)\rangle \rightarrow |1(1/2)\rangle \rightarrow |1(1)\rangle$. This sequence of states can be recorded as a path that is initially in the state $|0(0)\rangle$ and jumps to the state $|1(1/2)\rangle$ at time $s = 1/2$. The resulting path $(\{\nu\}, \{s\})$ is given by the sequences $\{\nu\} = \{0, 1\}$ and $\{s\} = \{0, 1/2\}$. Equivalently, this path can be expressed as a diagram, as seen in Fig. 1a.

Diagrams are a convenient way to think about the paths that describe adiabatic evolution. In our diagrams, an arrow denotes the evolution of the state as a function of time. If the arrow runs along a line, which denotes a particular energy level, then it implies that the system remains in that instantaneous eigenstate throughout that portion of the evolution. If the arrow connects two energy levels, then it implies that a transition, or jump, between energy levels occurs at that time. As level crossings are forbidden in our analysis, we do not change the relative spacing between the energy levels as the eigenvalue gap changes with time. Such diagrams are useful because they convey the intuition of what is being described by our more mathematical notation, which uses a tuple of sequences to describe the path. A similar diagrammatic notation is also presented by MMP [26].

In this work we classify paths by the number of jumps that they contain. As an example, consider the following evolution: $|0(0)\rangle \rightarrow |1(1/2)\rangle \rightarrow |2(3/4)\rangle \rightarrow |0(1)\rangle$. The corresponding path is parameterized by $\{\nu\} = \{0, 1, 2, 0\}$ and $\{s\} = \{0, 1/2, 3/4, 1\}$. This path jumps from one level to another three times and therefore we say that it contains three jumps. Similarly, the previous example contained one jump. The number of jumps present in a path is $|\{\nu\}| - 1$. If $(\{\nu\}, \{s\})$ is a q -jump path then we define $\{\nu\} := \{\nu_0, \nu_1, \dots, \nu_q\}$ and similarly $\{s\} := \{s_0, s_1, \dots, s_q\}$, where $\nu_0 = G$ and $s_0 = 0$. We choose ν_0 and s_0 in this fashion so that the p^{th} jump occurs at time s_p , which makes our results easier to read. A diagram of the path in this example is given in Fig. 1b.

Additionally, some paths describe adiabatic evolution and others do not. We call an n -jump path “adiabatic” if the final and initial state in the path are the same. For example, the path in Fig. 1b, is an adiabatic path whereas the path in Fig. 1a is not. We call any path that is not “adiabatic” a “non-adiabatic” path. We use the notation J_q to represent the set of all q -jump paths and J'_q to represent the set of all q -jump non-adiabatic paths. Non-adiabatic paths are crucially important to our work because the error in the adiabatic approximation is given by the sum over all non-adiabatic paths. This observation is especially relevant in light of the fact that the convergence problems noted by MMP [26] occur because their sums are not restricted to non-adiabatic paths.

We describe below, the sum over all such paths in the limit of continuous time.

Definition 2. Let $(\{\nu\}, \{s\})$ represent an arbitrary path in J_q , in the limit as $L \rightarrow \infty$. We then define the sum over the set all q -jump paths, J_q to be

$$\int d(J_q) := \sum_{\nu_1 \neq G} \sum_{\nu_2 \neq \nu_1} \cdots \sum_{\nu_{q-1} \neq \nu_q} \int_0^1 \cdots \int_0^{s_3} \int_0^{s_2} ds^q, \quad (11)$$

and similarly we define the sum over all q -jump non-adiabatic paths, J'_q , to be

$$\int d(J'_q) := \sum_{\nu_1 \neq G} \sum_{\nu_2 \neq \nu_1} \cdots \sum_{\nu_q \notin \{\nu_{q-1}, G\}} \int_0^1 \cdots \int_0^{s_3} \int_0^{s_2} ds^q, \quad (12)$$

where $ds^q = ds_1 \cdots ds_q$.

Now that we have introduced our notions of paths and the sum over all paths, we apply them to describe the time-evolution of quantum mechanical systems.

V. PATH INTEGRAL REPRESENTATION OF TIME-EVOLUTION

In this section we show that by taking the continuous limit of a discrete approach that is similar to that of Ambainis and Regev [25], we obtain a path integral representation of the time-evolution operator that is similar to that of Mackenzie, Marcotte and Paquette (MMP) [26], but without using perturbation theory. The convergence issues noted in [26] are ameliorated by noting that not all n -jump paths contribute to the error in the adiabatic approximation; however, the sum of all n -jump non-adiabatic paths does. Our goal in this section is to find an expression for the sum of all the non-adiabatic paths to the error in the adiabatic approximation. Upper and lower bounds for the contribution of these paths are proven in Sec. VII.

To simplify our proof, we choose the eigenstates of the Hamiltonian such that

$$\beta_{n,m}(s) := \langle \dot{n}(s) | m(s) \rangle = \begin{cases} 0, & \text{if } E_n(s) \equiv E_m(s) \\ \frac{\langle n(s) | \dot{H}(s) | m(s) \rangle}{E_n(s) - E_m(s)}, & \text{otherwise} \end{cases} \quad (13)$$

similarly to [25, 26]. Here $\beta_{n,m}(s)$ can be interpreted as a tunneling amplitude per unit of dimensionless time between the states $|n\rangle$ and $|m\rangle$. The choice $\beta_{n,n}(s) = 0$ corresponds to choosing the phases of the instantaneous eigenstates (which are dynamically irrelevant [31]) such that the geometric phase [18] is zero, which simplifies our proof dramatically. We use this choice of eigenstates to state in the following theorem, which is the main result of this section.

Theorem 3. *If $\|H(s)\|$ is bounded and $H(s)$ is piecewise-differentiable on the interval $[0, 1]$ then $U(1, 0)|G(0)\rangle$ is*

$$\begin{aligned} U(1, 0)|G(0)\rangle = & e^{-i \int_0^1 E_G(\xi) d\xi T} |G(1)\rangle + \int \left(e^{-i(\int_{s_1}^1 E_{\nu_1}(\xi) d\xi + \int_0^{s_1} E_G(\xi) d\xi)T} \beta_{\nu_1, G}(s_1) \right) |\nu_1(1)\rangle d(J_1) \\ & + \int \left(e^{-i(\int_{s_2}^1 E_{\nu_2}(\xi) d\xi + \int_{s_1}^{s_2} E_{\nu_1}(\xi) d\xi + \int_0^{s_1} E_G(\xi) d\xi)T} \beta_{\nu_2, \nu_1}(s_2) \beta_{\nu_1, G}(s_1) \right) |\nu_2(1)\rangle d(J_2) + \cdots, \end{aligned} \quad (14)$$

where $(\{\nu\}, \{s\})$ represents an arbitrary path in J_q and ν_i and s_i refer to particular element in the sequences, and β is defined in Eq. 13.

The series in Eq. 14 has an intuitive interpretation. It states that the time-evolution is a weighted sum over all paths that could describe the time-evolution of the quantum system, wherein the weight of a particular path is given by the products of the tunneling amplitudes multiplied by the dynamic phase picked up in that path. Adiabatic behavior occurs because in the limit of large T , the phases that are achieved by each path connecting $|G(0)\rangle$ and $|\nu(1)\rangle$ for $\nu \neq G$, vary considerably. Similarly, we expect that the contribution of the sum of all paths that begin in $|G(0)\rangle$ and end in $|G(1)\rangle$ to be suppressed given that the path contains two or more jumps. Hence we expect that interference effects will cause their contribution to become diminished in this limit, leaving only the zero jump path which links $|G(0)\rangle$ to $|G(1)\rangle$. This phase cancelation effect is demonstrated in Fig. 2.

We use these properties of the eigenstates to show the following lemma, which is used to evaluate the products of projectors that appear in our proof of Theorem 3.

Lemma 1. *Let $(\{\nu\}, \{s\})$ be a path in J_q with $\{s\} = \{0, j_1/L, \dots, j_q/L\}$ for an integer value of L then,*

$$\begin{aligned} \left(\prod_{j=j_{q-1}+1}^{j_q} P_{\nu_q} \left(\frac{j}{L} \right) \right) \cdots \left(\prod_{j=1}^{j_1} P_{\nu_1} \left(\frac{j}{L} \right) \right) = \\ |\nu_q(j_q/L)\rangle \langle \nu_1(0)| \left(\frac{\prod_{\ell=1}^{q-1} \beta_{\nu_{\ell+1}, \nu_\ell} \left(\frac{j_\ell}{L} \right)}{L^{q-1}} + O(j_q/L^{q+1}) \right), \end{aligned} \quad (15)$$

where each $P_{\nu_i}(j/L) = |\nu_i(j/L)\rangle \langle \nu_i(j/L)|$ projects onto the ray spanned by the state $|\nu_i(j/L)\rangle$.

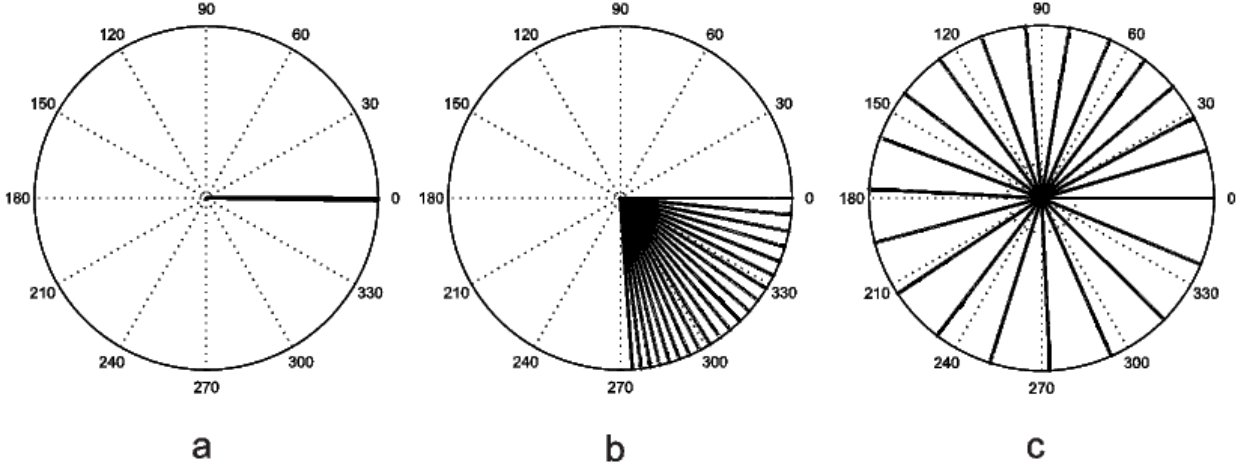


FIG. 2: This figure is a plot of the phase of states output by 21 one-jump paths that jump at time $s_1 = \{0, 1/20, 2/20, \dots, 1\}$ from the ground state and the first excited state for a Hamiltonian with $E_G(s) = 0$ and $E_1(s) = 1 + s$, for $T = 0.01$, $T = 1$ and $T = 4$ in a, b and c respectively. Here each solid line is a unit vector in \mathbb{C} that is oriented in the direction of $\exp(-i \int_{s_1}^1 E_1(\xi) - E_G(\xi) d\xi T)$. This plot shows that as T increases the vectors tend to become uniformly distributed throughout the unit circle, suggesting that the sum over all paths becomes less substantial as T increases.

Proof. We prove the lemma by performing induction over q , which we use to denote the number of jumps in the paths in question. Our first step is to prove the base case, $q = 1$. Using Taylor's theorem, and the fact that the eigenstates obey $\langle \dot{\nu}_\ell | \nu_\ell \rangle = 0$, it follows that for any ν_ℓ and s ,

$$\langle \nu_\ell(s + 1/L) | \nu_\ell(s) \rangle = 1 + O(1/L^2). \quad (16)$$

This implies that for any j_ℓ

$$\prod_{j=j_{\ell-1}+1}^{j_\ell} P_{\nu_\ell}(j/L) = |\nu_\ell(j_\ell/L)\rangle \langle \nu_\ell([j_{\ell-1} + 1]/L)| (1 + O([j_\ell - j_{\ell-1} - 1]/L^2)). \quad (17)$$

We demonstrate the base case, $q = 1$, of Eq. 15 by substituting $\ell = 1$ and $j_0 = 0$ into Eq. 17.

Now let us assume that the result of Lemma 1 applies for the case

$$\begin{aligned} & \left(\prod_{j=j_{q-2}+1}^{j_{q-1}} P_{\nu_{q-1}}\left(\frac{j}{L}\right) \right) \cdots \left(\prod_{j=1}^{j_1} P_{\nu_1}\left(\frac{j}{L}\right) \right) \\ &= |\nu_{q-1}(j_{q-1}/L)\rangle \langle \nu_1(0)| \left(\frac{\prod_{\ell=1}^{q-2} \beta_{\nu_{\ell+1}, \nu_\ell}(\frac{j_\ell}{L})}{L^{q-1}} + O(j_{q-1}/L^q) \right). \end{aligned} \quad (18)$$

By using Eq. 18 and Eq. 17 we find that the left hand side of Eq. 15 becomes

$$|\nu_q(j_q/L)\rangle \langle \nu_q([j_{q-1} + 1]/L)| \nu_{q-1}(j_{q-1}/L)\rangle \langle \nu_1(0)| \left(\frac{\prod_{\ell=1}^{q-2} \beta_{\nu_{\ell+1}, \nu_\ell}(\frac{j_\ell}{L})}{L^{q-1}} + O(j_q/L^q) \right). \quad (19)$$

Taylor's Theorem and the fact that, from the definition of a path, $\nu_q \neq \nu_{q-1}$ imply,

$$\langle \nu_q([j_{q-1} + 1]/L) | \nu_{q-1}(j_{q-1}/L) \rangle = \frac{1}{L} \langle \dot{\nu}_q(j_{q-1}/L) | \nu_{q-1}(j_{q-1}/L) \rangle + O(1/L^2). \quad (20)$$

$$= \frac{\beta_{\nu_q, \nu_{q-1}}(j_{q-1}/L)}{L} + O(1/L^2). \quad (21)$$

We substitute Eq. 21 into Eq. 19. The induction step on q follows, completing our proof of the lemma. \square

Lemma 1 provides an expression that can be used to evaluate products of projection operators. These results allow us to prove Theorem 3 by discretizing the time-evolution and examining the evolution in each time step in the basis formed by the instantaneous eigenstates of $H(s)$. We use projectors to transform the basis from the previous time step to the basis at the current time step. The resulting expression for the time-evolution is given as a weighted sum of products of projectors, which we then simplify by using Lemma 1. Our proof of Theorem 3 is given below.

Proof of Theorem 3. It follows from using the Trotter formula [32] for a fixed value of L that if $H(s)$ is piecewise differentiable then we can approximate $H(s)$ by a piecewise constant Hamiltonian to find

$$\begin{aligned} U(1,0) &= \prod_{j=0}^L \left(e^{-iH(j/L)T/L} \right) + O(1/L), \\ &= \prod_{j=0}^L \left(\sum_{\nu=0}^{N-1} P_{\nu}(j/L) e^{-iE_{\nu}(j/L)T/L} \right) + O(1/L), \end{aligned} \quad (22)$$

where for any $s \in [0,1]$ and $\nu \in \{0,1,\dots,N-1\}$, $|\nu(s)\rangle$ is an eigenstate of $H(s)$ and $P_{\nu}(s) := |\nu(s)\rangle\langle\nu(s)|$. We then expand Eq. 22 to find

$$U(1,0)|G(0)\rangle = \sum_{\nu_1=0}^{N-1} \cdots \sum_{\nu_{L-1}=0}^{N-1} \prod_{j=1}^L e^{-iE_{\nu_j}(j/L)T/L} P_{\nu_j}(j/L)|G(0)\rangle + O(1/L). \quad (23)$$

Equation 23 is a weighted sum over every possible path that can describe the evolution of the quantum system over the time interval. We group the paths in the sum by the number of jumps that comprise them and apply the result of Lemma 1 to find,

$$\begin{aligned} U(1,0)|G(0)\rangle &= e^{-i(\sum_{j=1}^L E_0(j/L)T/L)} |G(1)\rangle \\ &+ \sum_{\nu_1 \neq G} \sum_{j_1=1}^L \frac{e^{-i(\sum_{j=j_1}^L E_{\nu_1}(j/L) + \sum_{j=0}^{j_1-1} E_G(j/L))T/L} \beta_{\nu_1,G}(j_1/L)}{L} |\nu_1(1)\rangle \\ &+ \sum_{\nu_2 \neq \nu_1} \sum_{\nu_1 \neq G} \sum_{j_2=2}^L \sum_{j_1=1}^{j_2-1} \frac{e^{-i(\sum_{j=j_2}^L E_{\nu_1}(j/L) + \sum_{j=j_1}^{j_2-1} E_{\nu_1}(j/L) + \sum_{j=0}^{j_1-1} E_G(j/L))T/L}}{L^2} \\ &\quad \times \frac{\beta_{\nu_2,\nu_1}(j_2/L) \beta_{\nu_1,G}(j_1/L)}{L^2} |\nu_2(1)\rangle \\ &+ \cdots + O(1/L). \end{aligned} \quad (24)$$

Equation 24 is similar to that of Theorem 3, except instead of time-integrals we have discrete sums over time and we also have an $O(1/L)$ term that represents the error in simulating the adiabatic evolution using a piecewise constant Hamiltonian. We replace many of the sums in Eq. 24 by integrals, using the fact that

$$\sum_{j=s_1 L}^{s_2 L} f(j/L)/L = \int_{s_1}^{s_2} f(s)ds + O(1/L), \quad (25)$$

and eliminate the simulation error, by examining Eq. 24 in the limit as $L \rightarrow \infty$. This demonstrates that the sum of all one- and two-jump paths approach the path integrals in Theorem 3. The analogous results for paths that contain three or more jumps follow from the exact same reasoning, completing our proof of the theorem. \square

We have introduced our concept of path-integrals in this section, and have found an expression for the time-evolution operator in terms of these path integrals. Our next step is to approximate the contribution of the non-adiabatic paths in Eq. 14. This gives us the error in the adiabatic approximation. We use the results of the following section to find such a bound, which we give in Sec. VII.

VI. A RECURSIVE EXPRESSION FOR PATH INTEGRALS

Although an upper bound on the contribution one-jump non-adiabatic paths to the error in the adiabatic approximation can be directly calculated from Theorem 3, it is difficult to upper bound the contribution of paths that contain more than one jump using that strategy. To do so, we use the fact that every m -jump path is a concatenation of m one-jump paths. We use this property to find a recursion relation that allows us to express the contribution of an m -jump non-adiabatic path in terms of the contribution of an $(m-1)$ -jump non-adiabatic path (this fact follows from our definition of paths that we introduce in Sec. IV). Our error bounds then follow from summing a geometric series that results from unfolding these recursion relations.

Using the series in Eq. 14 and neglecting all paths that end in G (because such paths do not contribute to the approximation error), the error in the adiabatic approximation can be expressed as

$$((\mathbb{1} - P_G(1))U(1,0))|G(0)\rangle = \sum_{q=1}^{\infty} \left(\sum_{\nu_1 \neq G} \sum_{\nu_2 \neq \nu_1} \cdots \sum_{\nu_q \notin \{\nu_{q-1}, G\}} e^{-i \int_0^1 E_{\nu_q}(s) ds T} f_{\nu_q}(1) |\nu_q(1)\rangle \right), \quad (26)$$

where $\gamma_{\mu,\nu}(s)$ represents the eigenvalue gap between the instantaneous eigenstates $|\mu(s)\rangle$ and $|\nu(s)\rangle$ and f_{ν_q} represents the sum over all times as given by

$$\begin{aligned} f_{\nu_q}(1) &= \int_0^1 e^{-i \int_0^s \gamma_{\nu_{q-1}, \nu_q}(\xi) d\xi T} \beta_{\nu_q, \nu_{q-1}}(s_q) \\ &\times \int_0^{s_q} e^{-i \int_0^{s_{q-1}} \gamma_{\nu_{q-2}, \nu_{q-1}}(\xi) d\xi T} \beta_{\nu_{q-1}, \nu_{q-2}}(s_{q-1}) \times \cdots \\ &\times \int_0^{s_2} e^{-i \int_0^{s_1} \gamma_{G, \nu_1}(\xi) d\xi T} \beta_{\nu_1, G}(s_1) ds_1 \cdots ds_q. \end{aligned} \quad (27)$$

The integrals in Eq. 27 are expressed recursively as

$$\begin{aligned} f_{\nu_q}(x) &= \int_0^x e^{-i \int_0^s \gamma_{\nu_{q-1}, \nu_q}(\xi) d\xi T} \beta_{\nu_q, \nu_{q-1}}(s) f_{\nu_{q-1}}(s) ds, \\ f_{\nu_1}(x) &= \int_0^x e^{-i \int_0^s \gamma_{G, \nu_1}(\xi) d\xi T} \beta_{\nu_1, G}(s) ds. \end{aligned} \quad (28)$$

A similar recursive expansion to Eq. 28 is found by Sun in [27]. In the following section we find higher-order approximations to the error in the adiabatic approximation by using integration by parts to approximate the recursion relations in Eq. 28.

VII. APPROXIMATING PATH INTEGRALS

We have argued qualitatively that quantum interference effects are responsible for the emergence of adiabatic behavior in quantum systems. In this section we provide quantitative bounds for this phase cancellation effect. We first provide an example of how integration by parts can be used to approximate the phase cancellation effect for one-jump paths, and then use the same strategy to build an asymptotic expansion for the sum over all q -jump non-adiabatic paths in powers of T^{-1} , for every $q > 0$. We then sum these expressions over all values of q and obtain bounds for the approximation error. We then prove Theorem 2 using these bounds for the error.

A. Approximating Integrals Using Integration By Parts

In this subsection we provide an example that shows how integration by parts can be used to estimate the contribution of all one-jump non-adiabatic paths to the error in the adiabatic approximation. The purpose of this introduction is to provide a simple example of approximating path integrals that will help explain our approach for bounding the contribution of q -jump paths to the approximation error, which we address in the following subsection.

Integration by parts is a technique that is commonly used to approximate the integrals that appear in proofs of the adiabatic approximation [10, 26]. Unlike other methods, such as the method of steepest descent, integration by parts

gives an exact expression for the integral. Specifically, if $f : \mathbb{R} \mapsto \mathbb{C}$ and $g : \mathbb{R} \mapsto \mathbb{C}$ are differentiable functions and if $\partial_s g(s)$ is non-zero in the region of integration then [33]

$$\int_0^x f(s) e^{\lambda g(s)} ds = f(s) \left(\lambda \frac{\partial g(s)}{\partial s} \right)^{-1} e^{\lambda g(s)} \Big|_0^x - \int_0^x \left(\frac{\partial}{\partial s} \left[f(s) \left(\lambda \frac{\partial g(s)}{\partial s} \right)^{-1} \right] \right) e^{\lambda g(s)} ds. \quad (29)$$

Equation 29 gives the leading order behavior of the integral in the limit of large λ as well as an exact expression for the remainder term (which can also be approximated using integration by parts) given that $\partial_s g(s) \neq 0$ for all $s \in [0, x]$. For this reason, we assume that the minimum eigenvalue gap between any two non-degenerate eigenvectors is non-zero because, if it is ever zero, then integration by parts will fail to approximate Eq. 14. In such circumstances, an alternative method would have to be used to analyze the integral.

We approximate the contribution of all one-jump paths to the error in the adiabatic approximation as an example using integration by parts to bound a path integral. The contribution of all one-jump paths to the approximation error is

$$\int |\nu_1(1)\rangle \beta_{\nu_1, G}(s) e^{-i \int_0^s \gamma_{G, \nu_1}(\xi) d\xi T} d(J'_1) = \sum_{\nu_1 \neq G} |\nu_1(1)\rangle \int_0^1 \beta_{\nu_1, G}(s) e^{-i \int_0^s \gamma_{G, \nu_1}(\xi) d\xi T} ds. \quad (30)$$

We apply integration by parts to obtain,

$$\sum_{\nu_1 \neq G} |\nu_1(1)\rangle \left(\frac{\beta_{\nu_1, G}(s) e^{-i \int_0^s \gamma_{G, \nu_1}(\xi) d\xi T}}{-i \gamma_{G, \nu_1}(s) T} \Big|_{s=0}^1 - \int_0^1 \left(\frac{\partial}{\partial s} \left[\frac{\beta_{\nu_1, G}(s)}{-i \gamma_{G, \nu_1}(s) T} \right] \right) e^{-i \int_0^s \gamma_{G, \nu_1}(\xi) d\xi T} ds \right). \quad (31)$$

Using $\|\dot{H}\| := \max_s \|\dot{H}(s)\|$ and $\max_s \|\ddot{H}(s)\| \in O(1)$ we find that Eq. 31 has the upper bound

$$\left\| \int |\nu_1(1)\rangle \beta_{\nu_1, G}(s) e^{-i \int_0^s \gamma_{G, \nu_1}(\xi) d\xi T} d(J'_1) \right\| \leq 2 \frac{\|\dot{H}\|}{\min_{\nu_1, s} [\gamma_{\nu_1, G}(s)]^2 T} + O(1/T^2), \quad (32)$$

which is a commonly used estimate of the error in the adiabatic approximation. If the contribution of all non-adiabatic paths with two or more jumps is negligible, then Eq. 32 implies that the standard estimate for the approximation error holds [26].

Higher-order estimates for the contribution of the one-jump terms can be found by applying integration by parts to approximate the remainder term. The success of this strategy depends on the radius of convergence of the perturbation expansion. Unfortunately, the series has a zero-radius of convergence. This implies that arbitrarily high-order adiabatic theorems generated using our technique will only be useful in the limit as $T \rightarrow \infty$. This is because the repeated derivatives that occur from using integration by parts to obtain an $O(1/T^k)$ approximation to C_1 are $O(k!/T^k)$. The perturbation series for the one-jump path integrals will generally have a radius of convergence of zero as $k \rightarrow \infty$; suggesting that higher-order adiabatic approximations are often only useful in the limit of large T , as also noted in [13].

In the subsequent subsection we generalize the ideas presented in this subsection to find bounds for the approximation error that emerges due to *every* non-adiabatic path, as opposed to the results in this subsection, which considered only paths with one jump.

B. Error Bounds For The Adiabatic Approximation

Our strategy for bounding the error in the adiabatic approximation is to use integration by parts to find upper bounds for the oscillatory integrals present in Theorem 3. Our zeroth-order and first-order adiabatic approximations in Theorems 1 and 2, directly follow from this upper bound, which is given in the following Theorem.

Theorem 4. *Let $H : [0, 1] \mapsto \mathbb{C}^{N \times N}$ be a Hamiltonian that is differentiable three times and has a minimum eigenvalue gap between any two non-degenerate eigenvectors of $\gamma_{\min} > 0$. Then, for any $T \geq 0$, the error in the adiabatic approximation obeys*

$$\left\| (\mathbb{I} - P_G(1)) U(1, 0) |G(0)\rangle - \sum_{\nu \neq G} \frac{e^{-i \phi_\nu} |\nu(1)\rangle \langle \dot{\nu}(s_1) | G(s_1)\rangle e^{-i \int_0^{s_1} \gamma_{G, \nu}(\xi) d\xi T}}{-i \gamma_{G, \nu}(s_1) T} \Big|_{s_1=0}^1 \right\| \leq \frac{\mathbf{R}}{T^2}, \quad (33)$$

where $\phi_\nu = \int_0^1 E_\nu(\xi) d\xi T$ and \mathbf{R} is bounded above by

$$\mathbf{R} \leq \frac{2\|\dot{H}\| + \|\ddot{H}\|}{\gamma_{\min}^3} + \frac{25\|\dot{H}\|\|\ddot{H}\| + 16\|\dot{H}\|^2 + \|\ddot{H}\|^2}{\gamma_{\min}^4} + \frac{12\|\ddot{H}\|\|\dot{H}\|^2 + 118\|\dot{H}\|^3}{\gamma_{\min}^5} + \frac{36\|\dot{H}\|^4}{\gamma_{\min}^6} + T^2 \left[\left(1 + \frac{\Gamma}{\|\dot{H}\|T} \right) (e^{\Gamma/(\gamma_{\min}T)} - 1) - \frac{\Gamma}{\gamma_{\min}T} \right], \quad (34)$$

and $\Gamma = \left(\frac{6\|\dot{H}\|^3}{\gamma_{\min}^3} + \frac{\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^2} \right) + \frac{2\|\dot{H}\|^2}{\gamma_{\min}^2}$, $\|\dot{H}\|$, $\|\ddot{H}\|$ and $\|\ddot{H}\|$ refer to the maximum values of the norms of the first three derivatives of H and $\gamma_{G,\nu}(s) = E_G(s) - E_\nu(s)$.

This theorem is important because it provides upper and lower bounds that, except in pathological cases, converge to the observed approximation error in the limit of large T . As a result, these bounds are superior to the JRS bound [10] in the limit of large T . We compare these two bounds in greater detail in Sec. VIII.

The proof of Theorem 4 requires several steps, which we express in three lemmas. In order to abbreviate the statements of our Lemmas, we define the contribution of all j -jump non-adiabatic paths to the error in the adiabatic approximation to be C_j . Hence the error is

$$(I - P_G(1))U(1, 0)|G(0)\rangle = C_1 + C_2 + C_3 + \dots, \quad (35)$$

where for example,

$$C_1 := \int \left(e^{-i(\int_{s_1}^1 E_{\nu_1}(\xi) d\xi + \int_0^{s_1} E_G(\xi) d\xi)T} \beta_{\nu_1, G}(s_1) \right) |\nu_1(1)\rangle d(J'_1), \quad (36)$$

and

$$C_2 := \int \left(e^{-i(\int_{s_2}^1 E_{\nu_2}(\xi) d\xi + \int_{s_1}^{s_2} E_{\nu_1}(\xi) d\xi + \int_0^{s_1} E_G(\xi) d\xi)T} \beta_{\nu_2, \nu_1}(s_2) \beta_{\nu_1, G}(s_1) \right) |\nu_2(1)\rangle d(J'_2). \quad (37)$$

We have that $\nu_2 \neq G$ in our expression for C_2 because any such path is an adiabatic path and thus does not contribute to the error. The following lemma contains our upper bounds for the contribution of C_j for $j > 2$, which is an important step towards proving Theorem 4.

Lemma 2. *Let $H : [0, 1] \mapsto \mathbb{C}^{N \times N}$ be a Hamiltonian that is twice differentiable, then the following bound holds on the error in the adiabatic approximation:*

$$\|(I - P_G(1))U(1, 0)|G(0)\rangle - C_1\| \leq \|C_2\| + \left(1 + \frac{\Gamma}{\|\dot{H}\|T} \right) (e^{\Gamma/(\gamma_{\min}T)} - 1) - \frac{\Gamma}{\gamma_{\min}T}. \quad (38)$$

Here $\Gamma = \left(\frac{6\|\dot{H}\|^3}{\gamma_{\min}^3} + \frac{\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^2} \right) + \frac{2\|\dot{H}\|^2}{\gamma_{\min}^2}$ and $\|\dot{H}\|$ and $\|\ddot{H}\|$ refer to the maximum values of the norms of the first two derivatives of H and γ_{\min} is the minimum eigenvalue gap between any two non-degenerate eigenvalues.

Our proof of Lemma 2 uses integration by parts to approximate $\|C_j\|$ for $j > 2$. To do so we must provide upper bounds for the derivatives of β that appear due to using integration by parts to evaluate the integrals. The most straightforward way to approximate these derivatives is to differentiate Eq. 13 and use the triangle inequality on the result. However, this simple approach leads to extraneous factors of N , which appear because of the sums over ν_q . We use a trick involving the resolution of unity to provide upper bounds for the first two derivatives of β that do not depend on N . We state this result in the subsequent lemma, and will use it to prove Lemma 2.

Lemma 3. *Let $H(s)$ be twice differentiable on $s \in I = [0, 1]$ and let $s' \in I$, $\chi : J_{q-1} \mapsto \mathbb{C}$ be a bounded function, and for simplicity we define the argument of χ to be the path $\vec{\nu} := (\{\nu_j : j = 0, \dots, q-1\}, \{s_j : j = 0, \dots, q-1\})$ then*

$$\left\| \sum_{\nu_q, \nu_{q-1}, \dots, \nu_1} \dot{\beta}_{\nu_q, \nu_{q-1}}(s) \chi(\vec{\nu}) |\nu_q(s')\rangle \right\| \leq \left(\frac{4\|\dot{H}\|^2}{\gamma_{\min}^2} + \frac{\|\ddot{H}\|}{\gamma_{\min}} \right) \left\| \sum_{\nu_{q-1}, \nu_{q-2}, \dots, \nu_1} \chi(\vec{\nu}) |\nu_{q-1}(s)\rangle \right\| \quad (39)$$

and

$$\left\| \sum_{\nu_q, \nu_{q-1}, \dots, \nu_1} \ddot{\beta}_{\nu_q, \nu_{q-1}}(s) \chi(\vec{\nu}) |\nu_q(s')\rangle \right\| \leq \left(\frac{44\|\dot{H}\|^3}{\gamma_{\min}^3} + \frac{12\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^2} + \frac{\|\ddot{H}\|}{\gamma_{\min}} \right) \left\| \sum_{\nu_{q-1}, \nu_{q-2}, \dots, \nu_1} \chi(\vec{\nu}) |\nu_{q-1}(s)\rangle \right\|. \quad (40)$$

The function χ in these bounds is an arbitrary function, for example if we wish to find an upper bound for $\left\| \sum_{\nu_2, \nu_1} \dot{\beta}_{\nu_2, \nu_1}(s) \beta_{\nu_1, G}(s) |\nu_2(s')\rangle \right\|$, then we can use Eq. 40 to find the upper bound by choosing $\chi(\nu_1, s) = \beta_{\nu_1, G}(s)$. The proof of Lemma 3 is given in Appendix A.

We will now use the first result in Lemma 3 to prove Lemma 2.

Proof of Lemma 2. Using integration by parts on Eq. 28 yields

$$f_{\nu_q}(x) = \frac{f_{\nu_{q-1}}(s) e^{-i\kappa_{\nu_{q-1}, \nu_q}(s)T} \beta_{\nu_q, \nu_{q-1}}(s)}{-i\gamma_{\nu_{q-1}, \nu_q}(s)T} \Big|_0^x - \int_0^x \frac{\partial}{\partial s} \left(\frac{\beta_{\nu_q, \nu_{q-1}}(s) f_{\nu_{q-1}}(s)}{-i\gamma_{\nu_{q-1}, \nu_q}(s)T} \right) e^{-i\kappa_{\nu_{q-1}, \nu_q}(s)T} ds, \quad (41)$$

where $\kappa_{\nu_{q-1}, \nu_q}(s) = \int_0^s \gamma_{\nu_{q-1}, \nu_q}(\xi) d\xi$ and $f_{\nu_q}(x)$ represents the time-integrals in the sum over all q -jump non-adiabatic paths. Using Eq. 28 and the Fundamental Theorem of Calculus, we can simplify Eq. 41 by substituting

$$\partial_s f_{\nu_{q-1}}(s) = e^{-i\kappa_{\nu_{q-2}, \nu_{q-1}}(s)T} \beta_{\nu_{q-1}, \nu_{q-2}}(s) f_{\nu_{q-2}}(s). \quad (42)$$

We also have that for any $p > 0$, $f_p(0) = 0$. We then simplify Eq. 41 by using the recursion relations to express every instance of $f_{\nu_{q-1}}$ as an integral of $f_{\nu_{q-2}}$, to find

$$\begin{aligned} f_{\nu_q}(x) = & \left(\frac{e^{-i\kappa_{\nu_{q-1}, \nu_q}(x)T} \beta_{\nu_q, \nu_{q-1}}(x)}{-i\gamma_{\nu_{q-1}, \nu_q}(x)T} \int_0^x e^{-i\kappa_{\nu_{q-2}, \nu_{q-1}}(s_1)T} \beta_{\nu_{q-1}, \nu_{q-2}}(s_1) f_{\nu_{q-2}}(s_1) ds_1 \right. \\ & - \int_0^x \frac{\partial}{\partial s_1} \left(\frac{\beta_{\nu_q, \nu_{q-1}}(s_1)}{-i\gamma_{\nu_{q-1}, \nu_q}(s_1)T} \right) e^{-i\kappa_{\nu_{q-1}, \nu_q}(s_1)T} \int_0^{s_1} e^{-i\kappa_{\nu_{q-2}, \nu_{q-1}}(s_2)T} \beta_{\nu_{q-1}, \nu_{q-2}}(s_2) f_{\nu_{q-2}}(s_2) d^2s \\ & \left. - \int_0^x \left(\frac{\beta_{\nu_q, \nu_{q-1}}(s) \beta_{\nu_{q-1}, \nu_{q-2}}(s)}{-i\gamma_{\nu_{q-1}, \nu_q}(s)T} \right) e^{-i(\kappa_{\nu_{q-1}, \nu_q}(s) + \kappa_{\nu_{q-2}, \nu_{q-1}}(s))T} f_{\nu_{q-2}}(s) ds \right). \end{aligned} \quad (43)$$

We then express the sum over all values of ν_1, \dots, ν_q in Eq. 43 as

$$\sum_{\nu_q, \nu_{q-1}, \dots} f_{\nu_q}(x) |\nu_q(x)\rangle = g_q(x) - h_q(x), \quad (44)$$

where for any $q > 0$,

$$\begin{aligned} h_q(x) := & \sum_{\nu_q, \nu_{q-1}, \dots} |\nu_q(x)\rangle \int_0^x \int_0^{s_1} \left(\frac{\partial}{\partial s_1} \frac{\beta_{\nu_q, \nu_{q-1}}(s_1)}{-i\gamma_{\nu_{q-1}, \nu_q}(s_1)T} \right) e^{-i(\kappa_{\nu_{q-1}, \nu_q}(s_1) + \kappa_{\nu_{q-2}, \nu_{q-1}}(s_2))T} \\ & \times \beta_{\nu_{q-1}, \nu_{q-2}}(s_2) f_{\nu_{q-2}}(s_2) d^2s, \end{aligned} \quad (45)$$

where $d^2s = ds_2 ds_1$ and

$$\begin{aligned} g_q(x) := & \sum_{\nu_q, \nu_{q-1}, \dots} |\nu_q(x)\rangle \left(\int_0^x \frac{e^{-i(\kappa_{\nu_{q-1}, \nu_q}(x) + \kappa_{\nu_{q-2}, \nu_{q-1}}(s_1))T} \beta_{\nu_q, \nu_{q-1}}(x) \beta_{\nu_{q-1}, \nu_{q-2}}(s_1) f_{\nu_{q-2}}(s_1)}{-i\gamma_{\nu_{q-1}, \nu_q}(x)T} \right. \\ & \left. - \frac{e^{-i(\kappa_{\nu_{q-1}, \nu_q}(s_1) + \kappa_{\nu_{q-2}, \nu_{q-1}}(s_1))T} \beta_{\nu_q, \nu_{q-1}}(s_1) \beta_{\nu_{q-1}, \nu_{q-2}}(s_1) f_{\nu_{q-2}}(s_1)}{-i\gamma_{\nu_{q-1}, \nu_q}(x)T} ds_1 \right). \end{aligned} \quad (46)$$

We now will proceed to bound $\|g_q(x)\|$. To do so, we use the definition of β in Eq. 13 to find following bound, which is valid for any s', s'', s''' and s'''' .

$$\begin{aligned} & \left\| \sum_{\nu_q, \nu_{q-1}, \dots} |\nu_q(x)\rangle \beta_{\nu_q, \nu_{q-1}}(s') \beta_{\nu_{q-1}, \nu_{q-2}}(s'') f_{\nu_{q-2}}(s''') / \gamma_{\nu_{q-1}, \nu_q}(s''') \right\| \\ &= \left\| \frac{\sum_{\nu_q, \nu_{q-1}, \dots} |\nu_q(x)\rangle \langle \nu_q(s') | \dot{H}(s') | \nu_{q-1}(s') \rangle \langle \nu_{q-1}(s'') | \dot{H}(s'') | \nu_{q-2}(s''') \rangle f_{\nu_{q-2}}(s''')}{\gamma_{\nu_{q-1}, \nu_q}(s') \gamma_{\nu_{q-2}, \nu_{q-1}}(s'') \gamma_{\nu_{q-1}, \nu_q}(s''')} \right\| \\ &\leq \frac{\|\dot{H}\|^2}{\gamma_{\min}^3} \left\| \sum_{\nu_q} |\nu_q(x)\rangle \langle \nu_q(s') | \right\| \left\| \sum_{\nu_{q-1}} |\nu_{q-1}(s'')\rangle \langle \nu_{q-1}(s'') | \right\| \left\| \sum_{\nu_{q-2}, \dots} f_{\nu_{q-2}}(s''') |\nu_{q-2}(s''')\rangle \right\| \\ &\leq \frac{\|\dot{H}\|^2}{\gamma_{\min}^3} \left\| \sum_{\nu_{q-2}, \dots} f_{\nu_{q-2}}(s''') |\nu_{q-2}(s''')\rangle \right\|, \end{aligned} \quad (47)$$

where γ_{\min} is the minimum eigenvalue gap between any two non-degenerate eigenstates, for any $s \in [0, 1]$. We then have

$$\begin{aligned} \|g_q(x)\| &\leq \frac{2\|\dot{H}\|^2}{\gamma_{\min}^3 T} \int_0^x \left\| \sum_{\nu_{q-2}, \nu_{q-3}, \dots} f_{\nu_{q-2}}(s_1) |\nu_{q-2}(s_1)\rangle \right\| ds_1 \\ &\leq \frac{2\|\dot{H}\|^2}{\gamma_{\min}^3 T} \int_0^x (\|g_{q-2}(s_1)\| + \|h_{q-2}(s_1)\|) ds_1, \end{aligned} \quad (48)$$

where the last inequality follows from the triangle inequality. Using similar reasoning and our expression for the derivative of β in Eq. 39, we find that

$$\begin{aligned} \|h_q(x)\| &\leq \left(\frac{\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^3 T} + \frac{6\|\dot{H}\|^3}{\gamma_{\min}^4 T} \right) \int_0^x \int_0^{s_1} \left\| \sum_{\nu_{q-2}, \nu_{q-3}, \dots} f_{\nu_{q-2}}(s_2) |\nu_{q-2}(s_2)\rangle \right\| d^2 s \\ &\leq \left(\frac{\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^3 T} + \frac{6\|\dot{H}\|^3}{\gamma_{\min}^4 T} \right) \int_0^x \int_0^{s_1} (\|g_{q-2}(s_1)\| + \|h_{q-2}(s_1)\|) d^2 s \end{aligned} \quad (49)$$

We eliminate $\|g_{q-2}\|$ and $\|h_{q-2}\|$ in Eq. 48 and Eq. 49 iteratively by using Eq. 48 and Eq. 49 to provide upper bounds for $\|g_{q-2}\|$ and $\|h_{q-2}\|$ in terms of $\|g_{q-4}\|$ and $\|h_{q-4}\|$. This process is then repeated $\lfloor q/2 \rfloor$ times to eliminate $\|h_i\|$ and $\|g_i\|$ from our bounds for any i . In the case where q is even, no integrals will remain after performing this substitution $q/2$ times, whereas if q is odd, the remaining terms will have one integral left that has not been approximated using integration by parts.

Let us now consider the case where $q = 2m$. We expand $f_{\nu_{2m}}$, which is the time-integral part of C_{2m} , into a sum of 2^m terms by applying the recursion relations in Eq. 49 and Eq. 48. We then group these terms by the number of integrals present in a given term. It follows from combinatorics that, for any $\ell \in \{0, 1, \dots, m\}$, there are $\binom{m}{\ell}$ ways to obtain a term with $2m - \ell$ integrals and hence

$$\begin{aligned} &\|h_{2m}(x)\| + \|g_{2m}(x)\| \\ &\leq \sum_{\ell=0}^m \binom{m}{\ell} \int_0^1 \cdots \int_0^{s_{2m-\ell}} \left(\frac{2\|\dot{H}\|^2}{\gamma_{\min}^3 T} \right)^{m-\ell} \left(\frac{\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^3 T} + \frac{6\|\dot{H}\|^3}{\gamma_{\min}^4 T} \right)^\ell d^{2m-\ell} s \\ &\leq \sum_{\ell=0}^m \binom{m}{\ell} \frac{1}{(2m-\ell)!} \left(\frac{2\|\dot{H}\|^2}{\gamma_{\min}^3 T} \right)^{m-\ell} \left(\frac{\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^3 T} + \frac{6\|\dot{H}\|^3}{\gamma_{\min}^4 T} \right)^\ell \\ &\leq \frac{1}{m!(\gamma_{\min} T)^m} \sum_{\ell=0}^m \binom{m}{\ell} \left(\frac{2\|\dot{H}\|^2}{\gamma_{\min}^2} \right)^{m-\ell} \left(\frac{\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^2} + \frac{6\|\dot{H}\|^3}{\gamma_{\min}^3} \right)^\ell \\ &= \frac{\Gamma^m}{m!(\gamma_{\min} T)^m}, \end{aligned} \quad (50)$$

where $1/(2m - \ell)!$ is the volume of integration and the last line of Eq. 50 follows from the binomial theorem for

$$\Gamma := \left(\frac{2\|\dot{H}\|^2 + \|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^2} + \frac{6\|\dot{H}\|^3}{\gamma_{\min}^3} \right). \quad (51)$$

By summing Eq. 50 over all $m \geq 2$, we find that

$$\sum_{m=2}^{\infty} \|C_{2m}\| \leq e^{\Gamma/(\gamma_{\min} T)} - \frac{\Gamma}{\gamma_{\min} T} - 1. \quad (52)$$

Our next step is to find $\sum_{m=1}^{\infty} \|C_{2m+1}\|$. To do so, we use the fact that a path that contains $2m + 1$ jumps is the concatenation of a path that contains $2m$ jumps and a path that contains one jump. An upper bound on the contribution of non-adiabatic paths with $2m + 1$ jumps is found by taking the norm of the result for the $2m$ jumps

in Eq. 50 and multiplying it by an upper bound on the remaining one-jump path. Hence an upper bound on the contribution of these paths is

$$\begin{aligned}
\|C_{2m+1}\| &\leq \frac{\Gamma^m}{m!(\gamma_{\min}T)^m} \max_{s \in [0,1]} \left\| \sum_{\nu_1} \int_0^s \beta_{\nu_1,G}(s') e^{-i \int_0^{s'} \gamma_{G,\nu_1}(\xi) d\xi T} |\nu_1(s)\rangle d s' \right\|, \\
&\leq \frac{\Gamma^m}{m!(\gamma_{\min}T)^m} \max_s \left\| \sum_{\nu_1} |\nu_1(s)\rangle \left(\frac{\beta_{\nu_1,G}(s') e^{-i \kappa_{G,\nu_1}(s')T}}{-i \gamma_{\nu_1,G}(s')T} \right) \right\|_0^s \\
&\quad - \int_0^s \frac{\partial}{\partial s'} \left(\frac{\beta_{\nu_1,G}(s')}{-i \gamma_{\nu_1,G}(s')T} \right) e^{-i \kappa_{G,\nu_1}(s')T} d s' \left\|, \\
&\leq \frac{\Gamma^m}{m!(\gamma_{\min}T)^m} \frac{\Gamma}{\|\dot{H}\|T}.
\end{aligned} \tag{53}$$

By summing this result over $m = 1, 2, \dots$ we find that

$$\sum_{m=1}^{\infty} \|C_{2m+1}\| \leq \frac{\Gamma}{\|\dot{H}\|T} \left(e^{\Gamma/(\gamma_{\min}T)} - 1 \right). \tag{54}$$

Adding Eq. 52 to Eq. 54 yields

$$\sum_{j=3}^{\infty} \|C_j\| \leq \left(1 + \frac{\Gamma}{\|\dot{H}\|T} \right) \left(e^{\Gamma/(\gamma_{\min}T)} - 1 \right) - \frac{\Gamma}{\gamma_{\min}T}. \tag{55}$$

The proof of the Theorem then follows from the above results. By definition, the contribution of the one-jump and two-jump paths to the error in the adiabatic approximation are $\|C_1\|$ and $\|C_2\|$ respectively. Equation 38 then follows from the triangle inequality. \square

Now all that remains in our proof of the adiabatic approximation is to bound the one- and two-jump contributions to the error in the adiabatic approximation. The following lemma contains our bounds for $\|C_1\|$ and $\|C_2\|$.

Lemma 4. *Let $H : [0, 1] \mapsto \mathbb{C}^{N \times N}$ be differentiable three times then,*

$$\left\| C_1 - \sum_{\nu} \frac{e^{-i\phi_{\nu}} |\nu(1)\rangle \beta_{\nu,G}(s_1) e^{-i\kappa_{G,\nu}(s_1)T}}{-i\gamma_{G,\nu}(s_1)T} \right\|_0^1 \leq \mathbf{R}_0, \tag{56}$$

where $\phi_{\nu} = \int_0^1 E_{\nu}(\xi) d\xi T$,

$$\mathbf{R}_0 := \frac{2\|\ddot{H}\| + \|\ddot{H}\|}{\gamma_{\min}^3 T^2} + \frac{20\|\dot{H}\|\|\ddot{H}\| + 12\|\dot{H}\|^2}{\gamma_{\min}^4 T^2} + \frac{88\|\dot{H}\|^3}{\gamma_{\min}^5 T^2}, \tag{57}$$

and

$$\|C_2\| \leq \frac{\|\ddot{H}\|^2 + 4\|\dot{H}\|^2 + 5\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^4 T^2} + \frac{12\|\ddot{H}\|\|\dot{H}\|^2 + 30\|\dot{H}\|^3}{\gamma_{\min}^5 T^2} + \frac{36\|\dot{H}\|^4}{\gamma_{\min}^6 T^2}, \tag{58}$$

where $\|\dot{H}\| = \max_{s \in I} \|\dot{H}(s)\|$, $\|\ddot{H}\| = \max_{s \in I} \|\ddot{H}(s)\|$, $\|\ddot{H}\| = \max_{s \in I} \|\ddot{H}(s)\|$ and γ_{\min} is the minimum energy gap between any two non-degenerate energy levels.

The proof of Lemma 4 is similar to our proof of Lemma 2. The main difference between our proofs is that Eq. 43 is not used, and instead we apply integration by parts twice to guarantee that our approximations are correct to $O(1/T^2)$. We then use the triangle inequality to find upper bounds for the $O(1/T^2)$ terms and then use the triangle inequality to prove our error bounds. The details of this proof are given in Appendix B.

The proof of Theorem 4 follows from the results of Lemmas 2 and 4.

Proof of Theorem 4. The proof immediately follows by substituting the bounds for $\|C_1\|$ and $\|C_2\|$ in Lemma 4 into the result of Lemma 2. \square

The proof of our zeroth-order adiabatic theorem, given as Theorem 1, follows directly from this result. Similarly to the proof of Theorem 4, the proof of Theorem 2 follows from combining Lemma 2 and Lemma 4 and then showing that the terms that are proportional to T^{-2} are $O(\Delta_1/\gamma_{\min}T^2)$, where the timescale Δ_1 is given in Definition 1.

Proof of Theorem 2. Our asymptotic bounds are taken in the limit of small $\Delta_1/[\gamma_{\min}T^2]$ and large Δ_1 . From the assumption that Δ_1 is asymptotically greater than any constant, we find by substitution that each of the five terms that are in Eq. 56 are proportional to T^{-2} and their sum is $O(\Delta_1/[\gamma_{\min}T^2])$. This implies that

$$C_1 = \sum_{\nu} \frac{e^{-i \int_0^1 E_{\nu}(s) ds T} |\nu(1)\rangle \beta_{\nu,G}(s_1) e^{-i \kappa_{G,\nu}(s_1) T}}{-i \gamma_{G,\nu}(s_1) T} \Big|_0^1 + O(\Delta_1/[\gamma_{\min}T^2]). \quad (59)$$

It follows from substitution into Eq. 58 that $C_2 \in O(\Delta_1/[\gamma_{\min}T^2])$.

Similarly, we have that

$$\Gamma^2 \in O \left(\frac{\|\dot{H}\|^2 \|\ddot{H}\|^2}{\gamma_{\min}^4} + \frac{\|\dot{H}\|^4 \|\ddot{H}\|}{\gamma_{\min}^5} + \frac{\|\dot{H}\|^6}{\gamma_{\min}^6} \right). \quad (60)$$

The second term in the above equation is asymptotically greater than or equal to the first term if $\|\ddot{H}\| \in O(\|\dot{H}\|^2/\gamma_{\min})$. However, this condition implies that the third term in Eq. 60 is asymptotically dominant. Therefore,

$$\Gamma^2 \in O \left(\frac{\|\dot{H}\|^2 \|\ddot{H}\|^2}{\gamma_{\min}^4} + \frac{\|\dot{H}\|^6}{\gamma_{\min}^6} \right). \quad (61)$$

The first term in the above equation is dominant only if $\|\ddot{H}\|^2/\gamma_{\min} \in O(\|\ddot{H}\|)$. It follows from Eq. 61 that

$$\Gamma^2 \in O \left(\frac{\|\ddot{H}\|^3}{\gamma_{\min}^3} + \frac{\|\dot{H}\|^6}{\gamma_{\min}^6} \right) \in O(\gamma_{\min} \Delta_1). \quad (62)$$

From our assumption that $\Delta_1/(\gamma_{\min}T^2) \in o(1)$ we find,

$$\frac{\Gamma^2}{\gamma_{\min}^2 T^2} \in o(1), \quad (63)$$

which implies that

$$\frac{\Gamma}{\gamma_{\min} T} \in o(1). \quad (64)$$

This result is important because it allows us to replace $\exp(\Gamma/[\gamma_{\min}T])$ with a truncated Taylor series, without introducing any error terms that are exponentially large. Following this strategy, we find from using Taylor's theorem and Eq. 64 that the result in Eq. 55 is $O(\Delta_1/[\gamma_{\min}T^2])$, proving that the contribution of $\|C_3\|, \|C_4\|, \dots$ is $O(\Delta_1/[\gamma_{\min}T^2])$, which proves the theorem. \square

We then can state the proof of our zeroth-order adiabatic theorem as follows.

Proof of Theorem 1. The proof directly follows by applying the triangle inequality to the result of Theorem 2 and maximizing each term that is present. \square

Our upper bounds can be further tightened if we assume that the Hamiltonian has only two distinct energy levels for all s . This is because, in this case, all paths that contain an even number of jumps and start in the state $|G(0)\rangle$, must end in the state $|G(1)\rangle$ because of our choice of eigenvectors in Eq. 13. Therefore we neglect all paths that contain an even number of jumps. This result is stated in the following corollary.

Corollary 3. *Let $H(s)$ be a twice differentiable Hamiltonian with two distinct energy levels $E_G(s)$ and $E_1(s)$, where the eigenvectors corresponding to $E_1(s)$ are $(N-1)$ -fold degenerate. Given these assumptions, the error in the adiabatic approximation is bounded above and below by,*

$$\|(\mathbb{I} - P_G(1))U(1,0)|G(0)\rangle - C_1\| \leq \left(\frac{\Gamma}{\|\dot{H}\|T} \right) (e^{\Gamma/(\gamma_{\min}T)} - 1), \quad (65)$$

Proof. The proof of this corollary follows from our proof of the previous theorem. Although E_1 is $(N - 1)$ -fold degenerate, this system is an effective two-level system because our choice of eigenstates prevents any transition between two states with the same energy eigenvalue. Given this choice of eigenbasis, any path that has an even number of jumps is an adiabatic path, which are excluded from the sum.

Equation 54 then gives an upper bound on contribution to the error due to paths with more than one-jump because all paths with an even number of jumps do not contribute to the error in this case. Equation 65 follows from the definition of C_1 and the triangle inequality. \square

VIII. SEARCH HAMILTONIAN

In this section we apply our error bounds in Theorem 4 to the case of a Search Hamiltonian, and compare our results to the JRS bound [10]. In addition, we discuss a new method for choosing the integration time steps geometrically.

The Search Hamiltonian is a Hamiltonian whose ground state at $s = 0$ is an initial guess for the solution to a search problem and the ground state at $s = 1$ is chosen to be the state that is sought by the algorithm [7, 11]. If the Hamiltonian evolution obeys the adiabatic approximation then the initial guessed state will be transformed into the final solution state, thereby solving the search problem.

We denote the initial state $|\psi_0\rangle$ and the marked state $|\psi_1\rangle$, and the goal of the algorithm is to choose a Hamiltonian that generates an evolution that transforms the initial guessed state into the marked state in the adiabatic limit. We evolve our system according to a time-dependent Hamiltonian that is a linear interpolation between these two:

$$H(s) = (1 - s)H_0 + sH_1 = \mathbb{1} - (1 - s)|\psi_0\rangle\langle\psi_0| - s|\psi_1\rangle\langle\psi_1|.$$

Although the search Hamiltonian has N eigenstates, only $|1(s')\rangle$ has non-zero overlap with the ground state $|G(s)\rangle$ for any s and s' . Therefore the search Hamiltonian generates an effective two-dimensional evolution. The eigenvalue gap between the two relevant states, if $|\psi_0\rangle = \sum_{x=0}^{N-1} |x\rangle/\sqrt{N}$ and $|\psi_1\rangle$ is a state in the computational basis, is [11]

$$\gamma_{1,G}(s) = \sqrt{1 - 4\left(1 - \frac{1}{N}\right)s(1 - s)}. \quad (66)$$

If we define $\theta = \arccos(|\langle\psi_0|\psi_1\rangle|)$ then the instantaneous eigenvectors of the Hamiltonian are

$$|G(s)\rangle = \frac{\sin(\theta - \varphi(s))}{\sin \theta} |\psi_0\rangle + \frac{\sin \varphi(s)}{\sin \theta} |\psi_1\rangle, \quad (67)$$

and

$$|1(s)\rangle = \frac{\cos(\theta - \varphi(s))}{\sin \theta} |\psi_0\rangle - \frac{\cos \varphi(s)}{\sin \theta} |\psi_1\rangle, \quad (68)$$

where

$$\varphi(s) = \frac{1}{2} \arctan \left(\frac{s \sin(2\theta)}{(1 - s + s \cos(2\theta))} \right). \quad (69)$$

These expressions for the eigenvectors clearly show that the eigenvectors undergo rotation in the plane formed by the span of $|\psi_0\rangle$ and $|\psi_1\rangle$, where the rotation angle is given by $\varphi(s)$. Similarly, the rotation speed is given by $\dot{\varphi}(s)$. This implies that the eigenstates rotate more slowly near $s = 1/2$ than near $s = 0$ or $s = 1$, it makes sense to use an integration scheme that takes constant sized steps in φ rather than constant sized steps in s . This is because constant sized steps in s will lead to overly conservative time steps being taken near $s = 0$ or $s = 1$.

We find for the case where $\theta = \arccos(1/\sqrt{N})$ that

$$\varphi(s) = \frac{[\arctan(\sqrt{N-1}(2s-1)) + \arctan(\sqrt{N-1})]}{2}. \quad (70)$$

Interestingly, this parametrization has been used in the literature to interpolate adiabatic evolution in order to minimize the approximation error [11, 34]. To clarify, we use this interpolation for the timesteps taken in our numerical method, but use linear interpolation for the Hamiltonian.

We use this result to test our upper and lower bounds for the error in the adiabatic approximation by computing the time-evolution numerically using a modified version of Eq. 22 that uses evenly spaced steps in φ rather than s . The

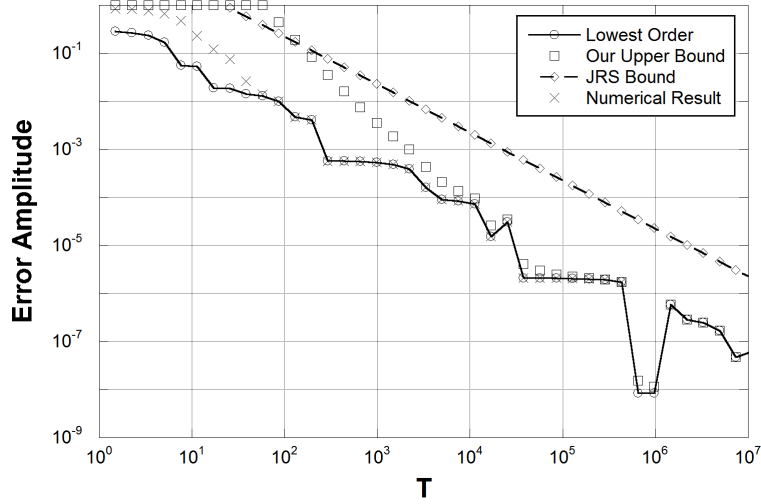


FIG. 3: This figure shows the excitation amplitude for a search Hamiltonian with 4 potential marked states with $|\psi_0\rangle = \sum_{v=0}^3 |v\rangle/\sqrt{4}$. The jaggedness of the figure appears because of rapid oscillations that occur in the limit of large T .

results were found by adaptively choosing L such that doubling its value resulted in less than a one percent change in the computed error in the adiabatic approximation. To test the results, we also compared the results of the simulation for small values of T to those found by solving the Schrödinger equation using an adaptive quadrature Runge-Kutta integrator using double precision data and found excellent agreement. However, we found that the Runge-Kutta integrator was unable to provide sufficient accuracy for simulations with large T and hence we used Eq. 22 in its place.

In Fig. 3 we use these expressions for the eigenvalue gap as well as the eigenvectors to evaluate our upper bound in Corollary 3, and compare them against the JRS bound [10] as well as our numerical data and our lowest order approximation to the error as given by Theorem 2. In all cases, our upper bound on the error is superior to the JRS bound, in the limit of large T . We also note that the JRS bound does not converge to the observed error in the limit of large T , whereas ours does. As expected from the fact that our bounds are higher-order than the JRS bound, our bounds are inferior for small values of T .

Fig. 4 demonstrates that our upper bound and lower bounds approach the result of Corollary 3 for large T . We do not plot our numerical data here because for $T > 100$ the numerical results are graphically indistinguishable from the result of Corollary 3 and the computation of the error in the approximation becomes prohibitively expensive for $T > 10^5$.

We cannot expect our bounds to be asymptotically tight for the search Hamiltonian because the leading order term, as predicted by Theorem 2 and Corollary 3, vanishes if

$$T = \frac{2n\pi}{\int_0^1 \gamma_{1,G}(s)ds}, \quad (71)$$

for any integer n . There are three steps in verifying this claim. First, we use Eq. 66 to see that $\gamma_{G,1}(0) = \gamma_{G,1}(1)$ for the search Hamiltonian. Second, we note that $\langle \dot{G}(1)|1(1)\rangle = \langle \dot{G}(0)|1(0)\rangle$ for the search Hamiltonian. Third and last, we substitute these results into Eq. 33 to find that the leading order term in both the upper and lower bounds vanish if T is chosen as in Eq. 71. As we only retained the exact expression for the $O(1/T)$ terms in the error in the adiabatic approximation, the higher order terms will not generically be correct. Therefore, we cannot expect our error bounds to be asymptotically tight because of the existence of such times.

On the other hand, if T is *not* chosen as in Eq. 71 then we can see that the bounds inferred from Eq. 33 approach each other as T increases. This follows because the norms of the derivatives of $H(s)$ are independent of T for the search Hamiltonian, and therefore the difference between the two bounds is $O(1/T^2)$. As the upper and lower bounds on the error approach each other, we know from the squeeze limit theorem that Eq. 33 and Eq. 33 approach the correct expression for the error at all points other than those in Eq. 71. Thus, our error bounds in Theorem 4 are not asymptotically tight in the limit of large T for the Search Hamiltonian; however, they are asymptotically tight for any Hamiltonian that is not explicitly a function of T , and for which the leading order contribution does not vanish at some $T \geq 0$.

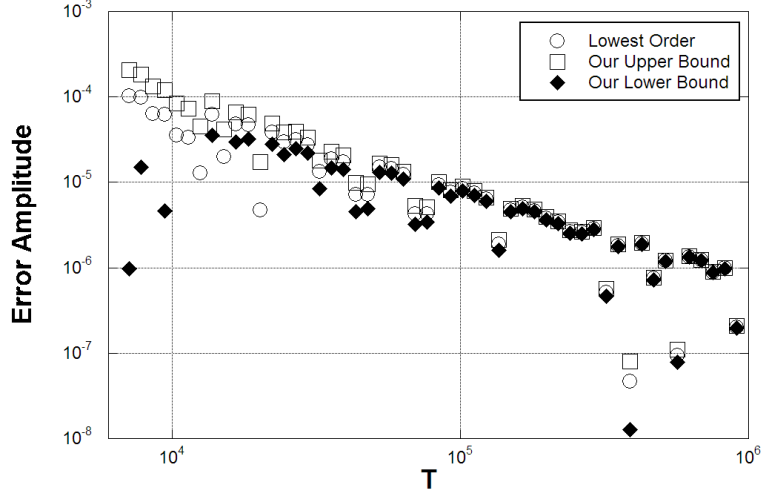


FIG. 4: This figure shows that our upper and lower bounds given by Theorem 4 for the error in the adiabatic approximation converge for a search Hamiltonian with 4 potential marked states, as $T \rightarrow \infty$.

IX. MARZLIN-SANDERS COUNTEREXAMPLE

The Marzlin–Sanders counterexample [14] is often placed at the center of the ongoing debate regarding the conditions that are needed to guarantee the validity of the adiabatic approximation [16, 17]. We examine two issues surrounding the counterexample, in this section. First, we look at the Marzlin–Sanders counterexample to benchmark our criteria for the validity of the adiabatic approximation. We do this to show that our sufficiency conditions prohibit the use of the adiabatic approximation for their Hamiltonian. Second, we examine the possibility that Hamiltonians that are related to the Marzlin–Sanders counterexample may actually behave adiabatically. In doing so, we provide proofs of Corollaries 1 and 2.

The counterexample Hamiltonian is of the form [14],

$$H_{MS}(s) = \left[\omega_0 \begin{pmatrix} \cos(2\pi s) \\ \sin(2\pi s) \\ 0 \end{pmatrix} + \frac{2\pi \sin(\omega_0 s T)}{T} \begin{pmatrix} -\sin(2\pi s) \cos(\omega_0 s T) \\ \cos(2\pi s) \cos(\omega_0 s T) \\ \sin(\omega_0 s T) \end{pmatrix} \right] \cdot [\sigma_x, \sigma_y, \sigma_z], \quad (72)$$

where σ_x , σ_y and σ_z are the Pauli operators. The eigenvalue gap for their Hamiltonian is

$$\gamma(s) = \sqrt{\omega_0^2 + (2\pi \sin(\omega_0 s T)/T)^2}. \quad (73)$$

The three most important quantities for determining the error using Eq. 38 are:

$$\|\partial_s H_{MS}(s)\| \in O(1), \quad (74)$$

$$\|\partial_s^2 H_{MS}(s)\| \in O(T), \quad (75)$$

$$\gamma_{\min} \in O(1). \quad (76)$$

By substituting these scalings into our upper bound in Theorem 4, we expect the error in the adiabatic approximation to be $\Theta(1)$ and not $o(1)$. This means that our results suggest, but do not imply, that the Marzlin–Sanders counterexample does not obey the adiabatic approximation. On the other hand, if the second derivative of the Marzlin–Sanders Hamiltonian were $o(T)$ instead of $\Omega(T)$, then the adiabatic approximation would hold for a sufficiently large value of T . We noted this fact in Corollary 2, which we prove below.

Proof of Corollary 2. Since $\|\dot{H}\| \in O(1)$ and $\|\ddot{H}\| \in o(T)$ it follows that $\Gamma \in o(T)$ and $\Gamma/T \in o(1)$. We have that

$$\left(1 + \frac{\Gamma}{\|\dot{H}\|T}\right) (e^{\Gamma/(\gamma_{\min}T)} - 1) - \frac{\Gamma}{\gamma_{\min}T} \in o(1). \quad (77)$$

Therefore Lemma 2 implies that

$$\|(\mathbb{1} - P_G(1))U(1,0)|G\rangle\| \leq C_1 + C_2 + o(1). \quad (78)$$

By applying similar reasoning, we find from Eq. 58 that $\|C_2\| \in o(1)$. As we made no assumptions about $\|\ddot{H}\|$ in the corollary, we cannot use Eq. 56 here. Instead, we see from Eq. 31 and Lemma 3 that $\|C_1\| \in o(1)$, given that $\|\ddot{H}\| \in o(T)$. Therefore if $\|\ddot{H}(s)\| \in o(T)$ then

$$\|(\mathbb{1} - P_G(1))U(1, 0)|G\rangle\| \in o(1). \quad (79)$$

□

The above argument shows that some quantum systems can obey the adiabatic approximation if the second derivatives of the Hamiltonian diverge sufficiently slowly. We also might imagine that if the derivatives were to diverge even slower then it could be possible that the standard criterion for the error in the adiabatic approximation holds for that evolution. This claim was made in Corollary 1, which we prove below using similar reasoning to that used in the proof of Corollary 2.

Proof of Corollary 1. Given that $\|\ddot{H}(s)\| \in o(\sqrt{T})$, Lemma 4 implies that $\Gamma/T \in o(1/\sqrt{T})$. Therefore we have that

$$\left(1 + \frac{\Gamma}{\|\dot{H}\|T}\right) (e^{\Gamma/(\gamma_{\min}T)} - 1) - \frac{\Gamma}{\gamma_{\min}T} \in o(1/T). \quad (80)$$

It follows from Eq. 58 that $\|C_2\| \in o(1/T)$. We then substitute $\|\ddot{H}(s)\| \in o(\sqrt{T})$ and $\|\ddot{H}(s)\| \in o(T)$ into Eq. 56 to find that

$$\|C_1\| \in O\left(\frac{\|\dot{H}(0)\|}{(\min_{\nu} \gamma_{G,\nu}^2(0))T} + \frac{\|\dot{H}(1)\|}{(\min_{\nu} \gamma_{G,\nu}^2(1))T}\right). \quad (81)$$

Hence it follows from Lemma 2 that

$$\|(\mathbb{1} - P_G(1))U(1, 0)|G\rangle\| \in O\left(\max_{s=0,1} \left[\frac{\|\dot{H}(s)\|}{(\min_{\nu} \gamma_{G,\nu}^2(s))^2T}\right]\right), \quad (82)$$

completing the proof. □

X. CONCLUSIONS

In this paper, we have shown that the traditional estimate of the error in the adiabatic approximation is valid if $H(s)$ is three-times differentiable and the norms of these derivatives are bounded above by a constant for all $T \geq 0$. We have also provided upper and lower bounds for the error in the adiabatic approximation that are asymptotically tight for many Hamiltonians. These bounds are highly relevant to applications in quantum information processing because they can be used to provide bounds for the resources needed to perform a quantum adiabatic computation using a particular adiabatic path.

We have also shown that some Hamiltonians that have second derivatives that diverge as $T \rightarrow \infty$ behave adiabatically. Interestingly, the Marzlin–Sanders counterexample Hamiltonian [14] marginally fails the criterion in that if the second derivative of their Hamiltonian diverged slightly slower as $T \rightarrow \infty$, then their proposed Hamiltonian would behave adiabatically in that limit. This implies that it is not necessary to forbid Hamiltonians that have second derivatives that diverge with T from the class of Hamiltonians that generate adiabatic evolution.

Our research also leaves open a number of avenues for further work. First, we have noted that for the search Hamiltonian, the leading order term in the expansion for the error in the adiabatic approximation vanishes at certain times, causing the error to transition from $O(1/T)$ to $O(1/T^2)$ at those times. This effect could be used to quadratically suppress the probability of a particular transition for an adiabatic evolution, and could lead to promising techniques for improving adiabatic quantum gates, adiabatic algorithms or quantum control protocols. This effect and its implications will be discussed elsewhere.

Second, a generalization of our bounds for unbounded Hamiltonians that act on an infinite-dimensional Hilbert space, or Hamiltonians for which $|G(0)\rangle$ is degenerate, would be an interesting contribution to the understanding of adiabatic phenomena. Similarly, our error bounds can be improved in situations where there are near degeneracies in the eigenvalues of the Hamiltonians, causing γ_{\min} to be small. In such circumstances it may be more appropriate to use integration by parts to approximate all jumps that occur across eigenstates that are separated by a large eigenvalue gap, and use a different approximation to handle transitions between nearly degenerate eigenstates.

Third, our analysis of the search Hamiltonian has shown that the error in the adiabatic approximation appears to fall below $2/3$ well before the regime where the adiabatic approximation holds. This suggests that bounds for the error in the adiabatic approximation are needed in the short-time regime in order to find the time complexity of adiabatic algorithms [13]. Such analysis has not yet been performed for adiabatic algorithms, implying that this remains an interesting avenue for further research.

Appendix A: Bounds for Derivatives of $\beta_{\mu,\nu}$

We prove our bounds on the derivatives of β that are presented in Lemma 3, in this appendix. These bounds allow us to estimate the contributions to the error in the adiabatic approximation that are incurred due the variation of the tunneling amplitude per unit time, β . Our proof of these bounds is given below.

Proof of Lemma 3. For convenience we omit the dependence on s of any function below in order to simplify our resulting expressions (IE: $\beta_{\mu,\nu} = \beta_{\mu,\nu}(s), \gamma_{\mu,\nu} = \gamma_{\mu,\nu}(s), \dots$). Then using this notation, we find by substituting the definition of $\beta_{\mu,\nu}$ into the LHS of Eq. 39 that $\left\| \sum_{\nu_q, \dots, \nu_1} \chi(\vec{\nu}) |\nu_q(s')\rangle \dot{\beta}_{\nu_q, \nu_{q-1}}(s) \right\|$ can be expressed as

$$\left\| \sum_{\nu_q, \dots, \nu_1} \chi(\vec{\nu}) |\nu_q(s')\rangle \left(\frac{\langle \dot{\nu}_q | \dot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}} + \frac{\langle \nu_q | \ddot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}} + \frac{\langle \nu_q | \dot{H} | \dot{\nu}_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}} - \frac{\dot{\gamma}_{\nu_q, \nu_{q-1}} \langle \nu_q | \dot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}^2} \right) \right\|. \quad (\text{A1})$$

The derivatives of every eigenstate are removed from Eq. A1 by employing a trick involving the resolution of unity. For a given eigenstate $|\nu\rangle = |\nu(s)\rangle$ we have that

$$\langle \dot{\nu} | = \sum_{m=0}^{N-1} \langle \dot{\nu} | m \rangle \langle m | = \sum_{m=0}^{N-1} \langle m | \beta_{m,\nu}. \quad (\text{A2})$$

If $\langle m | \nu \rangle = 0$, then it follows by differentiating that relationship that $\langle \dot{m} | \nu \rangle = -\langle m | \dot{\nu} \rangle$. Similarly, if $\nu = m$ then $\langle \dot{m} | \nu \rangle = 0 = -\langle m | \dot{\nu} \rangle$, by our choice of eigenbasis. Therefore it follows that

$$|\dot{\nu}\rangle = \sum_{m=0}^{N-1} |m\rangle \langle m | \dot{\nu} \rangle = - \sum_{m=0}^{N-1} |m\rangle \beta_{m,\nu}. \quad (\text{A3})$$

Then by using this property, Eq. A3, Eq. A2 and Eq. 13, the right hand side of Eq. A1 becomes

$$\left\| \sum_{\nu_q, \dots, \nu_1} \chi(\vec{\nu}) |\nu_q(s')\rangle \left(\sum_{m \neq \nu_q} \frac{\langle \nu_q | \dot{H} | m \rangle}{\gamma_{\nu_q, m}} \frac{\langle m | \dot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}} + \frac{\langle \nu_q | \ddot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}} - \sum_{m \neq \nu_{q-1}} \frac{\langle \nu_q | \dot{H} | m \rangle}{\gamma_{\nu_q, \nu_{q-1}}} \frac{\langle m | \dot{H} | \nu_{q-1} \rangle}{\gamma_{m, \nu_{q-1}}} - \frac{\dot{\gamma}_{\nu_q, \nu_{q-1}} \langle \nu_q | \dot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}^2} \right) \right\|. \quad (\text{A4})$$

Eq. Eq. A4 has now been re-written in a form where no derivatives of $|\nu_q\rangle$, $|\nu_{q-1}\rangle$ nor $|m\rangle$ are present. Furthermore the states $|\nu_q\rangle$ and $|m\rangle$ appear as parts of projection operators, which have unit norm. Then by using this property, the triangle inequality and lower bounding the eigenvalue gaps in the denominators by γ_{\min} we find that Eq. A4 is bounded above by

$$\left(\frac{2\|\dot{H}\|^2 + \|\dot{H}\| \max_{\nu_q, \nu_{q-1}} |\dot{\gamma}_{\nu_q, \nu_{q-1}}|}{\gamma_{\min}^2} + \frac{\|\ddot{H}\|}{\gamma_{\min}} \right) \left\| \sum_{\nu_{q-1}, \dots, \nu_1} \chi(\vec{\nu}) |\nu_{q-1}(s)\rangle \right\|. \quad (\text{A5})$$

Eq. Eq. 39 then follows by using $\max_{\nu_q, \nu_{q-1}} |\dot{\gamma}_{\nu_q, \nu_{q-1}}| \leq 2\|\dot{H}\|$.

By taking the derivative of the inside of the norm in Eq. A4, and by inserting the resolution of unity to eliminate all derivatives of eigenvectors, we find that

$$\left\| \sum_{\nu_q, \nu_{q-1}, \dots, \nu_1} \chi(\vec{\nu}) |\nu_q(s')\rangle \ddot{\beta}_{\nu_q, \nu_{q-1}}(s) \right\|$$

can be expressed as

$$\begin{aligned} & \left\| \sum_{\nu_q, \nu_{q-1}, \dots, \nu_1} \chi(\vec{\nu}) |\nu_q(s')\rangle \left(\frac{\partial}{\partial s} \left(\sum_{m \neq \nu_q} \frac{\beta_{\nu_q, m} \beta_{m, \nu_{q-1}} \gamma_{\nu_q, m}}{\gamma_{\nu_q, \nu_{q-1}}} - \sum_{m \neq \nu_q} \frac{\beta_{\nu_q, m} \beta_{m, \nu_{q-1}} \gamma_{m, \nu_{q-1}}}{\gamma_{\nu_q, \nu_{q-1}}} \right) \right. \right. \\ & + \sum_{m \neq \nu_q} \frac{\beta_{\nu_q, m} \langle m | \ddot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}} - \sum_{m \neq \nu_{q-1}} \frac{\beta_{m, \nu_{q-1}} \langle \nu_q | \ddot{H} | m \rangle}{\gamma_{\nu_q, \nu_{q-1}}} + \frac{\langle \nu_q | \ddot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}} - \frac{\dot{\gamma}_{\nu_q, \nu_{q-1}} \langle \nu_q | \ddot{H} | \nu_{q-1} \rangle}{\gamma_{\nu_q, \nu_{q-1}}^2} \\ & \left. \left. - \frac{\ddot{\gamma}_{\nu_q, \nu_{q-1}} \beta_{\nu_q, \nu_{q-1}} + \dot{\gamma}_{\nu_q, \nu_{q-1}} \dot{\beta}_{\nu_q, \nu_{q-1}}}{\gamma_{\nu_q, \nu_{q-1}}} + \frac{\dot{\gamma}_{\nu_q, \nu_{q-1}}^2 \beta_{\nu_q, \nu_{q-1}}}{\gamma_{\nu_q, \nu_{q-1}}^2} \right) \right\|. \end{aligned} \quad (\text{A6})$$

By applying the triangle inequality, using the definition of β in Eq. 13, the fact that for any x and y : $\gamma_{x,y}(s) \geq \gamma_{\min}$, $\|\beta_{x,y}(s)\| \leq \|\dot{H}\|/\gamma_{\min}^2$, $|\dot{\gamma}_{x,y}(s)| \leq 2\|\dot{H}\|$ and the fact that $\|\sum_m |m(s')\rangle \langle m(s)|\| = 1$ for any s and s' , we find that Eq. A6 is bounded above by

$$\begin{aligned} & \left\| \sum_{\nu_q, \nu_{q-1}, \dots, \nu_1} \chi(\vec{\nu}) |\nu_q(s')\rangle \left(\frac{\partial}{\partial s} \left(\sum_{m \neq \nu_q} \frac{\beta_{\nu_q, m} \beta_{m, \nu_{q-1}} \gamma_{\nu_q, m}}{\gamma_{\nu_q, \nu_{q-1}}} - \sum_{m \neq \nu_q} \frac{\beta_{\nu_q, m} \beta_{m, \nu_{q-1}} \gamma_{m, \nu_{q-1}}}{\gamma_{\nu_q, \nu_{q-1}}} \right) \right) \right\| \\ & + \left\| \sum_{\nu_q, \nu_{q-1}, \dots, \nu_1} \chi(\vec{\nu}) |\nu_q(s')\rangle \left(\frac{\ddot{\gamma}_{\nu_q, \nu_{q-1}} \beta_{\nu_q, \nu_{q-1}} + \dot{\gamma}_{\nu_q, \nu_{q-1}} \dot{\beta}_{\nu_q, \nu_{q-1}}}{\gamma_{\nu_q, \nu_{q-1}}} \right) \right\| + \frac{4\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^2} \\ & + \frac{\|\ddot{H}\|}{\gamma_{\min}} + \frac{4\|\dot{H}\|^3}{\gamma_{\min}}. \end{aligned} \quad (\text{A7})$$

Our final result in Eq. 40 is found by applying the triangle inequality to this, using Eq. 39 and the following result from [25],

$$\ddot{\gamma}_{\nu_q, \nu_{q-1}} \leq 2\|\ddot{H}\| + \frac{8\|\dot{H}\|^2}{\gamma_{\min}}. \quad (\text{A8})$$

We then find Eq. 40 by using the triangle inequality and equations Eq. A8 and Eq. 39 to upper bound Eq. A6. \square

Appendix B: Proof of Lemma 4

We present a proof of the upper bounds that are presented in Lemma 4 for the contribution to the error that results from the sum over all one- and two-jump paths. Our proof technique is similar to that used in the proof of Lemma 2. The most significant difference is that we have to apply integration by parts twice to approximate the sums of all one- and two-jump terms, which we denote C_1 and C_2 .

Proof of Lemma 4.

The contribution of the sum of all one-jump paths to the error in the adiabatic approximation is

$$C_1 = \sum_{\mu=1}^{N-1} |\mu(1)\rangle \int_0^1 \beta_{\mu, G}(s) e^{-i \int_0^s \gamma_{G, \mu}(\xi) d\xi} T ds. \quad (\text{B1})$$

We find the following upper bound on $\|C_1\|$ by applying integration by parts twice and the triangle inequality on Eq. B1

$$\begin{aligned} \|C_1\| & \leq \left\| \sum_{\nu_1} |\nu_1(1)\rangle \frac{\beta_{\nu_1, G}(s) e^{-i \kappa_{G, \nu_1}(s) T}}{-i \gamma_{G, \nu_1}(s) T} \right\|_0 \\ & + \left\| \sum_{\nu_1} |\nu_1(1)\rangle \frac{1}{-i \gamma_{G, \nu_1}(s) T} \left(\frac{\partial}{\partial s} \frac{\beta_{\nu_1, G}(s)}{-i \gamma_{G, \nu_1}(s) T} \right) e^{-i \kappa_{G, \nu_1}(s) T} \right\|_0 \\ & + \left\| \int_0^1 \sum_{\nu_1} |\nu_1(1)\rangle \left(\frac{\partial}{\partial s} \frac{1}{-i \gamma_{G, \nu_1}(s) T} \left(\frac{\partial}{\partial s} \frac{\beta_{\nu_1, G}(s)}{-i \gamma_{G, \nu_1}(s) T} \right) \right) e^{-i \kappa_{G, \nu_1}(s) T} ds \right\|. \end{aligned} \quad (\text{B2})$$

Using our bounds for the derivatives of β in Eq. 39 and Eq. 40, the bound $|\dot{\gamma}_{\nu_1, \nu}| \leq 2\|\dot{H}\|$ and the upper bound for $|\ddot{\gamma}|$ in Eq. A8 we find that

$$\|C_1\| \leq \left\| \sum_{\nu_1} |\nu_1(1)\rangle \frac{\beta_{\nu_1, G}(s) e^{-i\kappa_{G, \nu_1}(s)T}}{-i\gamma_{G, \nu_1}(s)T} \right\|_0^1 + \mathbf{R}_0. \quad (\text{B3})$$

where

$$\mathbf{R}_0 := \frac{2\|\ddot{H}\| + \|\ddot{H}\|}{\gamma_{\min}^3 T^2} + \frac{20\|\dot{H}\|\|\ddot{H}\| + 12\|\dot{H}\|^2}{\gamma_{\min}^4 T^2} + \frac{88\|\dot{H}\|^3}{\gamma_{\min}^5 T^2}. \quad (\text{B4})$$

Now we will apply similar reasoning to approximate C_2 , which is the contribution of all non-adiabatic two-jump paths to the error in the adiabatic approximation. This error can be expressed as,

$$\sum_{\nu_2 \neq 0} \sum_{\nu_1} |\nu_2(1)\rangle \int_0^1 \beta_{\nu_2, \nu_1}(s_1) e^{-i \int_0^{s_1} \gamma_{\nu_1, \nu_2}(\xi) d\xi T} \int_0^{s_1} \beta_{\nu_1, G}(s_2) e^{-i \int_0^{s_2} \gamma_{G, \nu_1}(\xi) d\xi T} ds_2 ds_1. \quad (\text{B5})$$

We use integration by parts once on the outer-most integral and the fundamental theorem of calculus to find

$$\begin{aligned} \sum_{\nu_2 \neq 0} \sum_{\nu_1} |\nu_2(1)\rangle & \left(\frac{\beta_{\nu_2, \nu_1}(1) e^{-i \int_0^1 \gamma_{\nu_1, \nu_2}(\xi) d\xi T}}{-i\gamma_{\nu_1, \nu_2}(1)T} \int_0^1 \beta_{\nu_1, G}(s_2) e^{-i \int_0^{s_2} \gamma_{G, \nu_1}(\xi) d\xi T} ds_2 \right. \\ & - \int_0^1 \left(\frac{\partial}{\partial s_1} \frac{\beta_{\nu_2, \nu_1}(s_1)}{-i\gamma_{\nu_1, \nu_2}(s_1)T} \right) e^{-i \int_0^{s_1} \gamma_{\nu_1, \nu_2}(\xi) d\xi T} \int_0^{s_1} \beta_{\nu_1, G}(s_2) e^{-i \int_0^{s_2} \gamma_{G, \nu_1}(\xi) d\xi T} ds_2 ds_1 \\ & \left. - \int_0^1 \frac{\beta_{\nu_2, \nu_1}(s_1) \beta_{\nu_1, G}(s_1) e^{-i \int_0^{s_1} \gamma_{G, \nu_1}(\xi) + \gamma_{\nu_1, \nu_2}(\xi) d\xi T}}{-i\gamma_{\nu_1, \nu_2}(s_1)T} ds_1 \right). \end{aligned} \quad (\text{B6})$$

We must have that $E_{\nu_2}(s) \neq E_G(s)$ to use integration by parts again on this expression, which is ensured by our assumption that the initial state is non-degenerate. Using integration by parts again to estimate the inner-most integrals in Eq. B6 we have that

$$\begin{aligned} \sum_{\nu_2, \nu} |\nu_2(1)\rangle & \left[\frac{\beta_{\nu_2, \nu_1}(1) e^{-i \int_0^1 \gamma_{\nu_1, \nu_2}(\xi) d\xi T}}{-i\gamma_{\nu_1, \nu_2}(1)T} \left(\frac{\beta_{\nu_1, G}(s_2) e^{-i \int_0^{s_2} \gamma_{G, \nu_1}(\xi) d\xi T}}{-i\gamma_{G, \nu_1}(s_2)T} \right) \right. \\ & - \int_0^1 \left(\frac{\partial}{\partial s_2} \frac{\beta_{\nu_1, G}(s_2)}{-i\gamma_{G, \nu_1}(s_2)T} \right) e^{-i \int_0^{s_2} \gamma_{G, \nu_1}(\xi) d\xi T} ds_2 \Big) \\ & - \int_0^1 \left(\frac{\partial}{\partial s_1} \frac{\beta_{\nu_2, \nu_1}(s_1)}{-i\gamma_{\nu_1, \nu_2}(s_1)T} \right) e^{-i \int_0^{s_1} \gamma_{\nu_1, \nu_2}(\xi) d\xi T} \left\{ \frac{\beta_{\nu_1, G}(s_2) e^{-i \int_0^{s_2} \gamma_{G, \nu_1}(\xi) d\xi T}}{-i\gamma_{G, \nu_1}(s_2)T} \right\}_{s_1} \\ & - \int_0^{s_1} \left(\frac{\partial}{\partial s_2} \frac{\beta_{\nu_1, G}(s_2)}{-i\gamma_{G, \nu_1}(s_2)T} \right) e^{-i \int_0^{s_2} \gamma_{G, \nu_1}(\xi) d\xi T} ds_2 \Big\} ds_1 \\ & - \frac{\beta_{\nu_2, \nu_1}(s_1) \beta_{\nu_1, G}(s_1) e^{-i \int_0^{s_1} \gamma_{G, \nu_1}(\xi) + \gamma_{\nu_1, \nu_2}(\xi) d\xi T}}{-\gamma_{\nu_1, \nu_2}(s_1)(\gamma_{G, \nu_1}(s_1) + \gamma_{\nu_1, \nu_2}(s_1))T} \Big|_0^1 \\ & \left. - \int_0^1 \left(\frac{\partial}{\partial s_1} \frac{\beta_{\nu_2, \nu_1}(s_1) \beta_{\nu_1, G}(s_1)}{-\gamma_{\nu_1, \nu_2}(s_1)(\gamma_{G, \nu_1}(s_1) + \gamma_{\nu_1, \nu_2}(s_1))T} \right) e^{-i \int_0^{s_1} \gamma_{G, \nu_1}(\xi) + \gamma_{\nu_1, \nu_2}(\xi) d\xi T} ds_1 \right]. \end{aligned} \quad (\text{B7})$$

Then using the triangle inequality, our upper bounds for the derivatives of β in Eq. 39 and Eq. 40, the bound $|\dot{\gamma}_{\nu_2, \nu}| \leq 2\|\dot{H}\|$ and the upper bound for $|\ddot{\gamma}|$ in Eq. A8 we find by substitution that

$$\|C_2\| \leq \frac{\|\ddot{H}\|^2 + 4\|\dot{H}\|^2 + 5\|\dot{H}\|\|\ddot{H}\|}{\gamma_{\min}^4 T^2} + \frac{12\|\ddot{H}\|\|\dot{H}\|^2 + 30\|\dot{H}\|^3}{\gamma_{\min}^5 T^2} + \frac{36\|\dot{H}\|^4}{\gamma_{\min}^6 T^2}. \quad (\text{B8})$$

□

Acknowledgments

We thank C. Dohotaru, J. Rashid, A. Reza khani, and B. C. Sanders for many helpful comments. This work was supported by Canada's Natural Sciences and Engineering Research Council (NSERC), the Canadian Institute for Advanced Research (CIFAR), Mathematics of Information Technology and Complex Systems (MITACS)'s Canadian Network of Centres of Excellence (NCE) program, the United States Army Research Office (USARO), and Quantum-Works.

-
- [1] J. Oreg, F. T. Hioe, and J. H. Eberly. *Phys. Rev. A*, **29**:690–697, 1984.
 - [2] J. R. Kuklinski, U. Gaubatz, F. T. Hioe, and K. Bergmann. *Phys. Rev. A*, **40**:67416744, 1989.
 - [3] F. T. Smith. *Phys. Rev.*, **179**:111–123, 1969.
 - [4] D. V. Averin. *Solid State Commun.*, **105**:659–664, 1998.
 - [5] N. S. Babcock, R. Stock, M. G. Raizen, and B. C. Sanders. *Can. J. Phys.*, **86**:549–555, 2008.
 - [6] D. Vager, B. Segev, and Y. B. Band. *Phys. Rev. A*, **72**:022325, 2005.
 - [7] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. *arXiv:quant-ph/0001106*, 2000.
 - [8] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev. In *Proc. 45th FOCS*, pages 42–51, 2004.
 - [9] J. Roland and N. J. Cerf. *Phys. Rev. A*, **68**:062311, 2003.
 - [10] S. Jansen, M. B. Ruskai, and R. Seiler. *J. Math. Phys.*, **48**:102111, 2007.
 - [11] J. Roland and N. J. Cerf. *Phys. Rev. A*, **65**:042308, 2002.
 - [12] E. Farhi, J. Goldstone, and S. Gutmann. *arXiv:quant-ph/0208135*, 2002.
 - [13] A. T. Reza khani, A. K. Pimachev, and D. A. Lidar. *arXiv:1008.0863*, 2010.
 - [14] K.-P. Marzlin and B. C. Sanders. *Phys. Rev. Lett*, **93**:160408, 2004.
 - [15] Z. Wu and H. Yang. *Phys. Rev. A*, **72**:012114, 2005.
 - [16] R. MacKenzie, A. Morin-Duchesne, H. Paquette, and J. Pinel. *Phys. Rev. A*, **76**(4):044102, 2007.
 - [17] M. S. Sarandy, L. A. Wu, and D. A. Lidar. *Quantum Inf. Process.*, **3**:331–349, 2004.
 - [18] M. V. Berry. *Royal Society of London Proceedings Series A*, **392**:45–57, 1984.
 - [19] T. Kato. *Journal of the Physical Society of Japan*, **5**:435–439, 1950.
 - [20] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, 1994.
 - [21] J. E. Avron and A. Elgart. *Commun. Math. Phys.*, **203**:445–463, 1999.
 - [22] J. E. Avron, R. Seiler, and L. G. Yaffe. *Comm. Math. Phys.*, **110**:33–49, 1987.
 - [23] B. W. Reichardt. In *Proc. 36th STOC*, pages 502–510, 2004.
 - [24] D. Teufel. *Adiabatic Perturbation Theory in Quantum Dynamics*, volume 1821 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2003.
 - [25] A. Ambainis and O. Regev. *arXiv:quant-ph/0411152*, 2004.
 - [26] R. MacKenzie, E. Marcotte, and H. Paquette. *Phys. Rev. A*, **73**:042104, 2006.
 - [27] C.-P. Sun. *J. Phys. A*, **21**:1595–1599, 1988.
 - [28] N. Wiener. *Proc. London Math. Soc.*, **22**:454–467, 1924.
 - [29] R. P. Feynman. *Rev. Mod. Phys.*, **20**:367–387, 1948.
 - [30] E. Farhi and S. Gutmann. *Annals of Physics*, **213**:182 – 203, 1992.
 - [31] L. E. Ballentine. *Quantum Mechanics, A Modern Development*. World Scientific, Singapore, 1998.
 - [32] H. Trotter. *Proc. Am. Math. Soc.*, **10**:545–551, 1959.
 - [33] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory*. Springer Verlag, New York, 1999.
 - [34] W. van Dam, M. Mosca, and U. Vazirani. In *Proc. 42th FOCS*, pages 279–287, 2001.